THE PHYSICAL SOCIETY

REPORTS ON PROGRESS IN PHYSICS

VOLUME XXI (1958)

EXECUTIVE EDITOR
A. C. STICKLAND

Published by
THE PHYSICAL SOCIETY
1 Lowther Gardens, Prince Consort Road
London S.W.7
1958

Printed in Great Britain by John Wright and Sons Ltd.
The Stonebridge Press, Bath Road, Bristol 4

CONTENTS

	PAGE
PREFACE	iv
Average Energy Expenditure per Ion Pair in Gases and Gas Mixtures. By J. M. VALENTINE and S. C. CURRAN	1
Absorption and Fluorescence Spectra of Ions in Crystals. By W. A. RUNCIMAN	30
Electromagnetic Excitation of Nuclei by Nuclear Projectiles and Electrons. By R. HUBY	59
Physics of Nerve Processes. By R. FATT	112
The Theory of Collective Excitations in Nuclei. By D. M. BRINK	144
Fourier Methods in Structure Analysis by Electron Diffraction. By J. M. COWLEY and A. L. G. REES	165
The Specific Heats of Metals at Low Temperatures. By D. H. PARKINSON	226
Mesonic Atoms. By D. WEST	271
Production and Measurement of High Temperatures. By W. LOCHTE-HOLTGREVEN	312

PREFACE

The *Reports on Progress in Physics* is under the general supervision of the Papers Committee of the Physical Society. The Members of the Papers Committee are :

Chairman : Professor R. W. Ditchburn.

Honorary Papers Secretary : Dr. H. H. Hopkins.

Ordinary Members : The President (Mr. J. A. Ratcliffe), Dr. M. Blackman, Professor W. E. Burcham, Dr. E. H. S. Burhop, Dr. A. H. Cooke, Professor S. Devons, Professor C. Domb, Mr. J. Dyson, Dr. W. Ehrenberg, Dr. B. H. Flowers, Professor F. C. Frank, Mr. D. W. Fry, Professor F. Llewellyn Jones, Dr. K. Mendelssohn, Professor R. E. Peierls, Mr. J. W. Perry, Dr. R. W. B. Stephens, Professor D. H. Wilkinson, Dr. D. A. Wright, Dr. C. G. Wynne, Dr. J. M. Ziman.

Reports published in the Volume are also being issued as individual articles, at prices depending on their length. Information may be obtained from the offices of the Society.

A. C. STICKLAND,
Secretary–Editor.

AVERAGE ENERGY EXPENDITURE PER ION PAIR IN GASES AND GAS MIXTURES

By J. M. VALENTINE† and S. C. CURRAN‡

† Western Regional Hospital Board, Regional Physics Department, Glasgow

‡ Atomic Weapons Research Establishment, Aldermaston, Berks.

CONTENTS

	PAGE
§ 1. Introduction	1
§ 2. Theoretical evaluation of w	2
§ 3. Electrons	3
3.1. External sources of radiation	4
3.1.1. Cloud chambers	4
3.1.2. X-ray and γ-ray beams	4
3.1.3. Electron beams	6
3.2. Internal sources of radiation	9
3.2.1. Proportional counter method	9
3.2.2. Ionization chamber and calorimetric measurements	11
3.2.3. Ionization chamber and chemical measurements	12
3.2.4. Standardization of radioactive isotopes	12
§ 4. Heavy particles	15
4.1. Alpha particles	15
4.1.1. Slow chambers	16
4.1.2. Fast chambers	18
4.1.3. Variation with energy	18
4.2. Protons	20
4.3. Heavy recoil particles and fission fragments	22
§ 5. Mixtures of gases	24
5.1. Low concentrations of one gas in another	24
5.2. Complete mixtures of two gases	25
References	27

Abstract. Theoretical and experimental work on the evaluation of w, the average energy expenditure per ion pair, in various gases and gas mixtures is reviewed. Ionization by electrons, α-particles, protons and heavy particles is discussed.

§ 1. Introduction

The phenomenon of ionization of gases is of fundamental importance in many fields of physics, both pure and applied. For example, many of our methods of detecting and measuring the energy of various types of radiation depend on the ionization, and consequently it is of prime importance that a sound theoretical and experimental knowledge of the process of energy loss of charged particles in gases should be secured. It might appear that the energy expenditure per ion pair produced could be related theoretically in a straightforward manner with several of the main parameters pertaining to the charged particle such as mass, charge and velocity. This is apparently not the case and much remains to be done, particularly in respect of theory, before we arrive at a really accurate and reliable presentation. A continuous and increasingly thorough attack has been made from the experimental side on the measurement

of energy expenditure and it is the authors' intention in the following paper to attempt to relate existing theory and present experimental knowledge and to assess where possible the overall situation, pointing at the same time to the various difficulties that arise and suggesting where further work would resolve some of the uncertainties. The subject has been reviewed in the past but the work continues unabated and it is felt that this review may not only present an up-to-date picture but help by drawing attention to various aspects which have not previously been considered in any detail.

§2. Theoretical Evaluation of w

The theoretical aspects of the penetration of atomic particles through matter have been reviewed by Bohr (1948). In particular the stopping power and ionization effects of particles in light substances are discussed and an expression is derived for the total number of collisions leading to ionization in the case of a primary particle. In order to obtain an expression for w, the amount of energy expended in the gas per ion pair produced, it is necessary to consider the total ionization, but it has not so far been possible to develop a general expression for the secondary ionization similar to that obtained for the primary ionization. It is possible, however, to deduce qualitatively how w will vary with type of ionizing particle and with particle energy and to calculate an approximate numerical value for w in the case of the simplest gases.

When an ionizing particle traverses a gas, part of the energy lost by the particle goes into processes other than ionization. From the point of view of ion yield this is ' wasted ' energy and so w must always be greater than the ionization potential I. The collisions made by a heavy particle may be subdivided into three groups: (i) excitation collisions in which ionization does not take place, (ii) light ionizing collisions in which the energy lost is of the order of I, and (iii) violent collisions in which high-speed secondary electrons are produced. According to Williams (1932), Bethe (1930) has shown that the relative numbers of excitation collisions and light ionization collisions are practically independent of the nature and velocity of the primary particle. This means that a constant amount of ' wasted ' energy is associated with the formation of an ion pair and so, neglecting violent collisions, w should be approximately constant. The relative numbers of high-speed secondary electrons produced in violent collisions is, of course, dependent on the energy of the primary particle, but violent collisions are infrequent and contribute little to the final total ionization. The high-speed secondary electrons also dissipate most of their energy ultimately in the form of light collisions and since, for such collisions, the ' wasted ' energy per ionizing collision is constant, it does not matter whether the ionizing particle is the primary particle or a fast secondary electron. Thus it is to be expected that w will prove to be more or less independent of the nature and velocity of the ionizing particle. This general conclusion will be seen to be in agreement with the experimental results.

A numerical estimate of the value for w is much more difficult to obtain. In an approximate calculation Bethe (1930) made a ' hydrogen-like ' approximation in which each molecular electron is treated as though it belonged to a hydrogen-like atom. Bagge (1937) made a more detailed calculation on the basis

of the Thomas–Fermi model of the atom. Both of these calculations give values of w of the right order of magnitude but they also predicted a positive correlation between w and I. This correlation is not supported by the experimental results. Fano (1946) calculated the values of w for atomic hydrogen and helium to be 36 and 38 ev respectively. In his calculation Fano took into account not only the screening of an electron by the inner electrons, but also the outer screening effect due to that part of the electronic cloud further away from the nucleus than the electron under consideration. Analysis of both these screening terms led Fano to a value for the effective atomic number. This was shown to remove the correlation between w and I. To obtain numerical values for w Fano used an approximate method in which he assumed that w is the same for both electrons and α-particles and is independent of particle energy in both cases. This removes the necessity of calculating directly the number of ion pairs produced by secondary electrons. An error in Fano's calculations has been pointed out by Erskine (1954 a). Using Fano's own data Erskine obtained the value of $w = 45$ ev for helium while noting that Fano's calculations refer to 10 kev electrons. The energy of α-particles having the same velocity is 73 Mev.

In this most recent calculation of w due to Erskine (1954 a), the ionization of helium, the simplest of the monatomic gases, by α-particles was considered. All the processes by which an α-particle can lose energy were considered and, using Born approximation formulae, detailed calculations were made of the cross sections for the excitation and ionization of helium atoms by α-particles and secondary processes. The final result was $w = 41{\cdot}1$ ev for 1 to 6 Mev α-particles. This is in very good agreement with the experimental value of 42·7 ev obtained by Jesse and Sadauskis (1953) for α-particles of energy 5·3 Mev in very pure helium. Experimental results for helium containing impurities to the level of only one part in a thousand indicate that the value for w drops to about 30 ev.

Although there is an adequate theoretical evaluation for w in the special case of very pure helium, the ionization of more complicated gases or gases which are not absolutely pure is not readily amenable to theoretical calculation. The experimental results have therefore to be considered without a theoretical framework which would lend shape to the discussion. Reviews of the experimental results have been made by Gray (1936, 1944), Binks (1936) and Compton and Allison (1936). Recently a very complete analysis of the numerical values obtained for w has been made by Binks (1954). Two factors of importance emerge—(i) the nature of the gas and (ii) the type and energy of the ionizing particle. A reasonable division can be made by considering first ionization by electrons (including x-rays and γ-rays) and secondly ionization by heavy particles.

§ 3. Electrons

Many of the experimental determinations of w for electrons have been carried out for air or nitrogen. This arises since air has long been used as a ' standard ' gas for the measurement of x-ray or γ-ray doses and a considerable variety of techniques have been used. A further sub-division can be effected by considering ' external ' and ' internal ' sources of radiation.

3.1. *External Sources of Radiation*

3.1.1. *Cloud chambers.*

A very direct method lies in the use of a Wilson cloud chamber. If an x-ray quantum of known energy suffers photo-electric absorption in the gas of the chamber a photo-electron of known energy E will be produced. Then if the expansion of the chamber is correctly made, the number of ions N produced by the electron is equal to the number of drops along its track and w is thus equal to E/N. Wilson (1923) and Meitner (1926) made measurements in this manner, but the values of w for air so obtained are not considered very reliable due to the difficulty in producing condensation on every ion without fog formation. This problem was examined in a more thorough investigation by Wilson and Dee (unpublished), who concluded that w for air lay between 32 and 34 ev for electrons of 9 kev initial energy.

A similar technique was used in a recent investigation by Frost and Nielson (1953) to determine specific probable ionization of electrons. In this case the electron momenta were found by the use of a magnetic field perpendicular to the direction of photography. The minimum ionization in ions per centimetre for dry gases under standard conditions was measured and theoretical minimum rates of energy loss divided by these values gave the following results for w:

Gas	Air	H_2	He	A
w (ev)	$31{\cdot}2 \pm 1{\cdot}5$	$37{\cdot}2 \pm 2{\cdot}0$	$26{\cdot}0 \pm 1{\cdot}6$	$27{\cdot}9 \pm 1{\cdot}5$

The method suffers from the obvious disadvantage that generally a mixture of gases or vapours is required for successful chamber operation.

3.1.2. *X-ray and γ-ray beams.*

If an x-ray beam of known intensity is partially absorbed in an ionization chamber and the amount of ionization measured, then a value for w can be determined. Two difficulties are inherent in this method: an estimation must be made of the extent to which the x-ray beam is absorbed and the intensity of the beam must be measured by a method independent of ionization. The usual method employed is a thermal one in which the x-ray beam is completely absorbed in a calorimeter and the heat produced is measured. In the early work with x-rays (up to 1926) the beams were rendered homogeneous by crystal reflections but the intensities available were very low and the necessary micro-calorimetric measurements were consequently rather uncertain. The first really consistent results were obtained by Kulenkampff (1926), who obtained beams of greater intensity by using the narrow spectral bands near the k_α lines of various elements and made the radiation homogeneous by means of filters. The value of w for air was found to be independent of energy in the range 22 to 6 kev with an absolute value of 35 ± 5 ev. The results of other investigators from 1926–29 who used similar methods are discussed by Compton and Allison (1936) and Binks (1936, 1954). The values of w_{air} obtained vary from 21 to 42·5 ev but no evidence was found in individual experiments of a variation of w with energy of the ionizing electrons. A different method of tackling the problem was reported by Steenbeck (1928), who measured the power of the x-ray beam by using a Geiger counter to count the actual number of absorption events in a

known thickness of gas. The work of Wilhelmy (1933) is also worthy of note because the radiation used was much softer (2–3 kev) than that employed by other workers. For such soft radiation no corrections are necessary for escaping energy. From the results of his relative measurements Wilhelmy concluded that w_{air} was increasing as the energy decreased in the range investigated.

In a series of investigations, Gaertner (1929 a, b, 1931 a, b) made measurements of w for various gases using a thermocouple to measure the energy of the beam. At the same time he used high-pressure ionization chambers so that practically all the energy of the x-ray beam was expended in the chamber and the amount of radiation escaping introduced only a small correction. This method of ' complete ' absorption was used by Crowther and Orton (1930, 1932) to measure the value of w for various gases relative to air as a standard, and their results were: N_2, 0·993; O_2, 1·103; CO_2, 1·035; H_2S, 1·41; A, 1·32. Since, in making relative measurements it is only necessary to have a constant source of x-rays and to measure saturation ion currents in the different gases, these relative measurements can generally be made with greater accuracy than absolute determinations.

Gaertner re-investigated the problem between 1933 and 1935 (Gaertner 1933, 1934/5, 1935) and obtained values of w for various gases relative to argon as reference gas. The method he employed was different from that used in his earlier experiments. Two ionization chambers were used in series, the second being filled with argon and the first with the gas under observation, and saturation ion currents were measured in both chambers for different pressures in the first chamber. From these measurements the ratio w_{gas}/w_A can be obtained. By keeping the pressure in the chambers low Gaertner avoided the criticism which was made of his earlier work, of incomplete saturation in the current measurements. Finally, to convert the relative measurements into absolute values, Gaertner measured w for argon absolutely and obtained $28{\cdot}4 \pm 0{\cdot}5$ ev—a value in good agreement with his previous result of $28{\cdot}8 \pm 1{\cdot}0$ ev found by using a different method. The values of w obtained by Gaertner are shown in table 1, column (1).

Table 1. w (ev)

Gas	Gaertner (1)	Crowther and Orton (2)	Crowther and Orton (3)
A	28·4†	28·4	26·7
N_2	35·82	37·8	35·4
O_2	31·64	34·0	31·9
Air	35·15	37·5	35·2†
CH_4	29·78	—	—
CO_2	33·0	36·3	34·0

† Reference gas.

For the sake of comparison the value of $w_A = 28{\cdot}4$ ev can be inserted into the ratio w_{air}/w_A obtained by Crowther and Orton giving a value of $w_{air} = 37{\cdot}5$ ev. This value can then be used to find absolute values of w from the ratios w_{air}/w_{gas} found by Crowther and Orton. These values are shown in column (2) of the table and are seen to differ considerably from the values obtained by Gaertner. On the other hand, if Gaertner's derived value of $w_{air} = 35{\cdot}2$ ev is used with Crowther and Orton's data, the figures in column (3) are obtained. Only the

value of w_A is seen to differ appreciably from Gaertner's results. From this it would appear that the ratio w_{air}/w_A obtained by Crowther and Orton is inaccurate and this point was investigated very carefully by Gaertner, who claimed to have proved conclusively that his ratio is the correct one. In his most recent work Gaertner quotes the mean wavelength of the x-rays with which he was working as 1·025 Å, corresponding to an electron energy of about 12 kev.

Similar experiments have been carried out recently by Laughlin, Beattie, Henderson and Harvey (1953) in a determination of the roentgen using x-ray energies of 400 kev and 22·5 Mev. They used a calorimetric method to compare the experimental values of the energy flux per roentgen in an absorbing metal cylinder with theoretical values. These results have been recalculated by Binks and Laughlin (quoted by Binks 1954) to yield values for w in air of $31{\cdot}0 \pm 1{\cdot}0$ and $31{\cdot}9 \pm 1{\cdot}6$ for the 400 kev and 22·5 Mev energies respectively. Other values by Laughlin (communicated to Binks 1954) are $w_A = 23$ ev and $w_{H_2} = 36{\cdot}0$ ev for 17·5 Mev electrons. Binks comments on the large correction applied for energy lost in the ion chamber in these experiments and this may account for the unusually low value for argon.

The development of thermistors capable of detecting a temperature change of less than 10^{-5} deg has been exploited by Johns, Berner and Cormack (1955) in their calorimetric measurements of the value of w_{air} for γ-rays from a unit containing 1000 curies of ^{60}Co. They claim that their experimental arrangement enabled them to measure changes in temperature of the order of 10^{-3} deg to an accuracy of better than 1%. To obtain a value of 33 ev for w_{air} they assumed a theoretical value for the mass stopping power for air.

3.1.3. *Electron beams.*

As an alternative to the use of x-rays electrons may be produced directly by a hot filament. These are accelerated through a known potential and the total charge of the beam is measured together with the ionization current produced when the beam is completely absorbed in a gas. From these measurements a value of w for the gas can be found. The electrons must necessarily be accelerated in an evacuated vessel which must then be connected to the gas-filled ionization chamber. Two methods of connecting the chambers have been used: (i) For electrons of energy greater than 10 kev, a thin window can be placed between the evacuated electron gun and the ionization chamber. The disadvantage of this method is that the energy of the electrons at the moment of entering the chamber is somewhat uncertain due to losses in the window and also the production of slow secondary electrons. (ii) Very slow primary electrons can be absorbed in a gas at very low pressure and so for primary electron energies of less than about 3 kev the ionization chamber can be directly connected to the electron gun by means of a long narrow tube. A good vacuum can be maintained in the accelerating chamber by means of fast pumps. This method is, however, also open to criticism concerning the precise energy of the electrons as they enter the ionization chamber since ions may be formed towards the high-pressure end of the capillary tube. If these ions are not collected the observed current will be too small with respect to the energy of the electrons as measured by the accelerating voltage and a spuriously high value for w would result.

A further difficulty inherent in the electron beam method is the necessity of producing a homogeneous electron beam in which the electron has the full energy corresponding to the potential difference through which it has been accelerated.

An early measurement by this method is due to Johnson (1917), who used slow electrons of the order of 200 ev and a direct connection between the ionization and electron accelerating chambers. He measured the total ionization produced in nitrogen, hydrogen, helium and oxygen and found that the values for w were similar for the four gases. The numerical values obtained were $w = 36$ ev for N_2 and O_2 and 41 ev for H_2 and He. Anslow (1925) performed a similar experiment in the energy range 200–1000 ev but the results obtained (about 5 ev for air) are not consistent with the process of ionization; a possible reason for this discrepancy may lie in the inhomogeneity of the beam as this factor was not investigated.

Lehmann and Osgood (1927) made a thorough investigation into the attainment of homogeneous electron beams and a specially designed Faraday cylinder was used to prevent loss of primary or secondary electrons in the electron current measurement. The value for w_{air} obtained was 45 ev for electron energy range 200–1000 ev. In a further paper Lehmann (1927) reports the values $w_{He} = 31$, $w_A = 33$, $w_{H_2} = 37$ and $w_{N_2} = 45$ ev. The same value of 45 ev was obtained by Schmitz (1928), who used a similar method with slightly faster electrons (1–9 kev). Buchmann (1928) connected the evacuated chamber to the ionization chamber by means of a celluloid window and obtained a much lower value for w_{air}, viz. 31 ± 3 ev for the energy range 4–13 kv.

One of the most exhaustive and accurate researches using this method is due to Eisl (1929). Velocity selection of electrons from the source was achieved by means of a magnetic field and homogeneity to $\pm \frac{1}{2}\%$ is claimed. A celluloid window $0{\cdot}1\mu$ thick separated the gun from the ionization chamber. The velocity distribution of the electron beam after it had passed through the window was investigated by the method of magnetic separation with photographic recording so that the initial energy of the electrons was known with certainty. Various other sources of error were investigated by Eisl and the final results of $w_{air} = 32{\cdot}6$ ev is claimed to be correct to within $\pm 1\frac{1}{2}\%$. A further conclusion drawn by Eisl was that w_{air} is constant for ionizing electrons in the range 9–59 kev.

In order to extend the energy range to lower value a similar investigation was carried out by Pigge (1934), who employed a capillary tube between the electron source and the chamber. To keep the pressure in the ionization chamber low the dimensions were large and the electrons were wound up into helical paths by means of a magnetic field. Electron and ion currents were measured with a valve galvanometer and the beam was shown to be homogeneous within narrow limits. The greatest source of uncertainty in this method lies in the formation, in the connecting tube, of ions which are not included in the current measurement. Pigge performed experiments to determine the magnitude of this error and estimated the necessary correction. He obtained 26 values for w in the energy range 3·05 to 0·315 kev. Of these, 16 are in the range 3·05 to 1·00 kev and are approximately constant with a mean value of 36 ev and a mean square error for the group of 2·5%. The remaining 10 values vary from 40 to 43·5 ev and these results contain an estimated correction of from 5 to 12%. Considering the difficulties

inherent in working with beams of very slow electrons, these results may be open to doubt. The experiments were performed in nitrogen instead of air to avoid oxidation of the hot filament.

Gerbes (1935) discussed Eisl's work and commented on the necessity of applying a small correction to the results to allow for an extra acceleration of the primary electrons by the field applied in the current measurement. He carried out an experiment to determine the size of this correction which, when applied to Eisl's results, gave a very small variation of w with particle energy. By converting Pigge's results to refer to air and combining them with those of Eisl, Gerbes obtained an empirical formula

$$w = 31{\cdot}62 + 5{\cdot}21/(E-I)^{1/2} \pm 0{\cdot}08 \text{ ev}$$

for the variation of w with particle energy from 0·3 to 60 kev. (E = initial energy of electrons, I = ionization potential of the gas.) This formula is largely based on the very low energy results of Pigge and so cannot be considered to have a very sound experimental basis.

This work was continued by Gerbes (1937) to determine the values of w for other gases with an apparatus similar to that of Eisl but incorporating several minor improvements. The energy range investigated was 10 to 40 kev and the results obtained by Gerbes are contained in table 2. Two values of w are given, these corresponding to the values obtained for the particular energies of 10 and 40 kev.

Table 2

Gas	w_{gas}/w_{air}	w (ev)	
		10 kev	40 kev
Air	1·00	33·4	32·4
N_2	1·037	34·6	33·6
O_2	0·914	30·5	29·6
CO_2	0·964	32·2	31·2
C_3H_8	0·722	23·6	23·7
SO_2	0·935	30·4	30·5
A	0·790	27·3	25·2
Ne	0·872	29·1	

These experiments show a similar variation of w_{air} with energy to that predicted by Gerbes' formula, but it is interesting to note that Gerbes obtains an even bigger variation in the case of argon while all the more recent results point to the constancy of w_A. The observations were made in the form w_{gas}/w_{air} and converted to absolute values by measuring w_{air}. (Variations in the ratio w_{gas}/w_{air} with particle energy were also found by Gerbes, but these are not included in the table.)

A further investigation of the low energy region was made by Freund (1935), who described an ingenious method for overcoming the two major difficulties (inaccurate ionization current and inhomogeneity of the electron beam) caused by the direct connection of the two chambers. By employing a three-slit system with variable potentials between the slits Freund was able to measure the number of ions formed by a group of electrons known to have the same velocity inside the ionization chamber. He concluded that for electron energies between 100 and

300 ev, w had an almost constant value about 30 ev, and w increased steadily for energies below 100 ev. This method was further refined by Breunig (1936), who used ultra-violet radiations to produce electrons by a photoelectric process and so obtained a more constant beam intensity. Breunig found, in agreement with Freund, a constant value for w (in this case 34 ev) in the range 160–70 ev and an increase of w below 70 ev.

3.2. *Internal Sources of Radiation*

The difficulty of introducing a beam of slow electrons into an ionization chamber is immediately overcome if the source of electrons is a radioactive material placed inside the chamber itself. The ready availability of beta emitting isotopes led to the introduction of several new methods, particularly suited to the determination of relative values in the form $w_{\text{gas}_1}/w_{\text{gas}_2}$ since only two ionization current measurements are required. Many of the more recent experimental results have been obtained in this form. Another feature of this work has been the use of argon rather than air as a ' standard ' gas, this stemming largely from the use of argon as a filling gas in counters and ionization chambers.

3.2.1. *Proportional counter method.*

The use of proportional counting tubes for the determination of w was first described by Curran, Cockroft and Insch (1950). If a source of ionizing radiation is contained in a tube, the count-rate n sec^{-1} can be accurately determined by operating the tube as a proportional counter. Further, if the tube is operated as an ionization chamber by reducing the applied voltage below the value for the onset of gas multiplication, the saturation ion current i can be measured. If the average energy of the radiations, $\bar{E}$, is known and e represents the electronic charge, then w can be determined from the equation

$$i = ne\bar{E}/w.$$

In practice, using electrons as ionizing agents, a count-rate which can be measured accurately is too small to give an accurately measurable ionization current. This difficulty can be overcome by counting a known fraction f ($\sim 1\%$) of the source used to produce the ionization current. When an isotope with a short half life is used (for example ^{64}Cu) f can be determined by allowing the source to decay for a known time. In the case of gaseous sources such as ^{3}H and ^{37}A, f can be found by sharing with known volumes.

In the original work Curran *et al.* made measurements with the help of the K x-rays from ^{64}Cu and the β-particles of ^{3}H on air, N_2, O_2, A, CO_2, CH_4 and C_2H_4. The actual numerical values are now considered suspect due to inaccuracies in the sharing technique and other difficulties. The technique was improved by Valentine (1952), who used proportional tubes corrected for end effect (Cockroft and Curran 1951). In his experiments the electrons and x-rays from the K-capture decay of ^{37}A and the β-rays from ^{3}H were employed. Three sets of measurements were made: absolute values using ^{37}A, relative values using ^{37}A, and relative values using ^{3}H. The absolute values in electron volts were

$$w_{\text{A}} = 27{\cdot}0\ ;\ w_{\text{He}} = 32{\cdot}2\ ;\ w_{\text{H}_2} = 37{\cdot}2\ ;\ \text{and}\ w_{\text{N}_2} = 35{\cdot}2.$$

To evaluate a small correction factor in this calculation it was necessary to assume a value for the Auger electron yield in the K-shell of chlorine. This was taken as 94%, but this figure has been criticized by Binks (1954), who suggested a value of 89% from a curve published by Broyles, Thomas and Haynes (1953). Recent experimental determinations of the Auger electron yield for argon have been carried out by Bertolini, Bisi and Zappa (1953), Heintze (1955) and Harrison, Crawford and Hopkins (1955). The numerical values obtained are respectively 92%, 87% and 92%, suggesting that 94% may be only a little high. The values of w for different gases are not equally sensitive to the figure used for the Auger yield; w_A and w_{N_2} are reduced by about 1% for a change from 94% to 89% but w_{H_2} and w_{He} fall by 5%. Other gases were measured relative to argon as the standard gas. Substituting the value of 27·0 ev for w_A the relative values in electron volts obtained were:

	w_{He}	w_{H_2}	w_{N_2}	w_{air}	w_{O_2}	w_{CH_4}
^{37}A	33·0	38·4	35·4			
^{3}H		38·0	36·0	35·0	32·3	29·5

No significant differences were found in the values of w obtained by the different methods and the final results of these studies, reported by Valentine and Curran (1952) and claimed to be correct to $\pm 0{\cdot}5$ ev, are shown in table 3.

Table 3

Gas	w (ev)	w_{gas}/w_A
A	27·0	1·00
He	32·5	1·20
H_2	38·0	1·41
N_2	35·8	1·32
Air	35·0	1·29
O_2	32·2	1·19
CH_4	29·5	1·09

Relative values of w which may be compared with those in table 3 were obtained from an experiment by Jesse and Sadauskis (1955). A cylindrical ionization chamber with a collecting electrode in the form of a wire grid was used and the central area of this grid was coated with the beta emitting isotopes under investigation. These were ^{63}Ni, which was electroplated on the wires, and ^{3}H in the form of tritiated polystyrene which was applied by a brush to the grid. The ionization currents produced in the chamber when it contained each of the gases were measured with the help of a vibrating reed electrometer used as a null indicator. The results in the form of w_{gas}/w_A are shown in table 6. This experiment was also concerned with the values of w obtained by ionization with α-particles and is discussed from that point of view in § 4.1.3.

A similar type of experiment to determine w_{air} has been described by Failla and Rossi (1952). A solid source containing ^{35}S was deposited on a very thin foil which was then suspended in the centre of a large ionization chamber and the saturation ion current in air measured. The rate of emission of β-particles from the source was determined using the same chamber after it has been completely evacuated. The value of w obtained, 32·5 ev, does not agree with the result of later work reported from the same laboratory (Failla 1956) and described in the

next section. Values of w, relative to air, were also measured (Gross 1954); these are: CH_4, 0·820; C_2H_2, 0·780; C_2H_4, 0·778; CO_2, 0·968; N_2, 1·034 and O_2, 0·914.

The proportional counter technique is a powerful method for the study of the variation of w with particle energy. If pulses from electrons or X-rays of known energy are recorded in a proportional counter, then a curve of pulse height against quantum energy can be constructed. If this curve is a straight line then w, for the gas contained in the counter, is constant. This method was used by Curran, Angus and Cockroft (1949), who showed that w_A was constant in the range 10 to 40 kev. The energy range has since been extended down to the L radiations of ^{37}A at about 240 ev (Pontecorvo 1950, Valentine 1952) and up to more than 1 Mev (see Fulbright 1955). From about 3 kev upwards it has been shown that w_A is constant within an experimental accuracy of about 2%. There is no evidence to show that w_A varies in the low energy region down to 240 ev but the rather large statistical widths of the very low energy peaks make the limits of accuracy rather greater. The method is also applicable to other gases provided they are suitable for proportional counter operation. Curran *et al.* (1949) showed w for nitrogen was constant in the range 5 to 25 kev and the low energy range using ^{37}A was studied by Valentine (1952). This investigation showed that w_N, did not increase by more than 5% for the range 46·5 kev to 2·8 kev and that there is not more than a 5% increase in the range 2·8 kev to 240 ev. This gives a possible variation of 10% between 46·5 kev and 240 ev. It is thought that w does not vary by as much as this but the rather poorer energy resolution of a nitrogen-filled counter makes it difficult to reduce the upper limit. This result is compatible with the low energy work of Freund (1935) and Breunig (1936) but does not agree with the observations of Pigge (1934).

3.2.2. *Ionization chamber and calorimetric measurements.*

Beta emitting radioactive isotopes have been used in a different way to obtain an absolute value for w. An extrapolation ionization chamber has been described by Failla (1956) and essentially it consists of a parallel plate ionization chamber with a variable air gap. Failla has described its use in eliminating the influence of an air cavity on the flux of ionizing particles when making ionization chamber measurements.

If a homogeneous medium contains a uniform concentration of a β-emitting isotope then application of the Bragg–Gray principle yields the following relation

$$3{\cdot}7 \times 10^{10} C\bar{E} = N_{air} w_{air} S_{air}^{s},$$

where C = concentration of radioactive material in curies g^{-1} of medium, $\bar{E}$ = average energy/disintegration in ev, N_{air} = number of ion pairs $(g_{air})^{-1}\,sec^{-1}$, w_{air} = ev/ion pair for β's in air, S_{air}^{s} = average true mass stopping power of the medium with respect to that of air for the β-particle flux in the medium.

Failla used this equation to measure w by determination of the other four factors. One electrode of his extrapolation chamber was a solution of the radioisotope and in the work carried out so far this has been ^{35}S. The ionization chamber measurements yielded a value of N_{air} and the product $C \times \bar{E}$ was determined by a calorimetric means method for about 450 mc of ^{35}S in the form of

sodium sulphate. A calculated value for S_{air}^{s} was adopted but it is intended that this factor should be observed experimentally for water vapour in later experiments. A preliminary value of $w_{air} = 35$ ev is quoted by Failla for this work.

Two studies by Hardwick (unpublished) and Dorfman (1954) of the ionization by tritium β-particles have yielded values of w relative to w_A. These are

	He	H_2	N_2	Air	O_2	CH_4
Hardwick	1·16	1·41	1·33	1·28	1·18	1·05
Dorfman	1·28	1·45			1·18	

3.2.3. *Ionization chamber and chemical measurements.*

Values of w were obtained for very high energy electrons by Weiss and Bernstein (1955, 1956) using a combination of a ferrous sulphate dosimeter and an ionization chamber. Measurements were made with X-rays from a 2 Mev Van de Graaff generator giving a spectrum with an average energy of about 1 Mev. Measurements of ionization currents in a parallel plate chamber with various gas fillings together with theoretical values for stopping powers give relative values for w. These were translated into absolute values by assessing the energy of the beam with a ferrous sulphate dosimeter. In this measurement a previously determined value of G, the oxidation yield of the ferrous sulphate solution, was used. The values of w obtained were: air, 33·9; H_2, 36·3; He, 40·3; Ne, 35·3; A, 25·8; Kr, 24·7; Xe, 22·0; O_2, 31·2; N_2, 34·6; C_2H_4, 26·4; CH_4, 26·8 and CO_2, 32·6 ev. The estimated errors on these figures vary within the range $\pm 0{\cdot}4$ to $\pm 0{\cdot}8$ ev.

3.2.4. *Standardization of radioactive isotopes.*

In a recent review article Emery (1956) shows how a value of w for air may be deduced from measurements on the disintegration of certain radioactive isotopes. Emery gives the following table, which is deduced from data collected by the National Physical Laboratory Advisory Committee on Radioactive Standards reported by Perry (1953) and American and Canadian results quoted by Seliger and Schwebel (1954).

Table 4

Isotope	Mean initial energy of electrons (Mev)	Range of values for w (ev) 4π β-counter	β–γ coincidence
^{24}Na	1·205	32·1–32·6 (4)	33·1–33·9 (3)
^{60}Co	0·592	32·1–34·2 (5)	34·3–34·8 (5)
^{198}Au	0·136	32·5–33·8 (4)	33·8–35·0 (4)
^{131}I	0·135	32·5–34·4 (6)	33·0–33·9 (2)
^{32}P	0·695	34·0–34·4 (7)	

Figures in parentheses give the number of independent determinations considered.

The results discussed above for ionization by electrons are summarized in two tables. Table 5 gives the values of w_{air} and the variation of w_{air} with energy while table 6 displays the results for other gases with relative values in the form w_{gas}/w_A shown in brackets.

Table 5

Author	Date	Method	Energy of electrons	w	Variation with E	Notes
Wilson and Dee	~1933	Cloud chamber	9 kev	32–34	—	—
Frost and Nielson	1953	Cloud chamber	High	$31{\cdot}2 \pm 1{\cdot}5$	—	Includes theoretical value for minimum rates of energy loss
Kulenkampff	1926	x-rays using micro-calorimetry	6–22 kev	35 ± 5	Independent	—
Gaertner	1929–35	x-rays ('total' absorption) + microcalorimetry	12 kev	35·2	—	—
Laughlin *et al.*	1954	x-rays + microcalorimetry	400 kev, 22·5 Mev	$31{\cdot}0 \pm 1{\cdot}0$, $31{\cdot}9 \pm 1{\cdot}6$	—	Includes theoretical values plus large correction factor, may be suspect
Johns *et al.*	1955	γ-rays from 1000 curie ^{60}Co unit, calorimetry with thermistors	1·1 Mev	33	—	Includes theoretical value for mass stopping power
Johnson	1917	Electron beam—capillary connection	~200 ev	36	—	—
Lehmann and Osgood	1927	Electron beam—capillary connection	200–1000 ev	45	None stated	—
Schmitz	1928	Electron beam—capillary connection	1–9 kev	45	None stated	—
Buchmann	1928	Electron beam—celluloid window	4–13 kev	31 ± 3	None stated	—
Eisl	1929	Electron beam—celluloid window	9–59 kev	32·6	None	—

Table 5 (*cont.*)

Author	Date	Method	Energy of electrons	w	Variation with E	Notes
Pigge	1934	Electron beam—capillary connection	3·05–0·314 kev	36–43·5	Yes	—
Gerbes	1935	Combination of results of Eisl and Pigge	—	—	$31{\cdot}62+\frac{5{\cdot}21}{(E-I)^{1/2}}$	—
Gerbes	1937	Electron beam—celluloid window	10–40 kev	33·4–32·4	As above	—
Freund	1935	Electron beam with slit system	100–300 ev	~30	None	Experiments performed to investigate variation with energy rather than establish absolute value
Breunig	1936	Ultra-violet production of electrons + Freund method	70–160 ev	~34	None	
Curran *et al.*	1950	Proportional counter	5–25 kev	—	None	
Valentine	1952	Proportional counter	2·82 and 5 kev	$35{\cdot}0\pm0{\cdot}5$	—	^{37}A and ^{3}H inside counter
Valentine	1952	Proportional counter	240 ev to 46·5 kev	—	Small or none	L-capture from ^{37}A to γ-rays from RaI
Failla and Rossi	1952	Ionization chamber	~70 kev	32·5	—	β-rays from ^{35}S
Failla	1956	Extrapolation ionization chamber plus calorimetry	~70 kev	35	—	450 mc of ^{35}S
Emery	1956	Standardization of radioactive isotopes	Various	32–35	None apparent	^{24}Na, ^{60}Co, ^{198}Au, ^{131}I, ^{32}P
Weiss and Bernstein	1955	Ionization chamber and ferrous sulphate dosimeter	~1 Mev	33·9	—	x-rays from 2 Mv Van de Graaff machine. Includes theoretical values for stopping powers

Table 6

	Gaertner (1929–35)		Crowther and Orton (1930, 1932)	Laughlin *et al.* (1954)		Johnson (1917)	Lehmann (1927)	
Gas	w (ev)	w_{gas}/w_A	w_{gas}/w_A	w (ev)	w_{gas}/w_A	w (ev)	w (ev)	w_{gas}/w_A
He						41·0	31·0	(0·94)
Ne								
A	28·4	(1·00)	(1·00)	23·0	(1·00)		33·0	(1·00)
Kr								
Xe								
H_2				36·0	(1·57)	41·0	37·0	(1·12)
Air	35·15	(1·24)	(1·32)	31·5	(1·37)		45·0	(1·36)
N_2	35·82	(1·26)	(1·33)			36·0	45·0	(1·36)
O_2	31·64	(1·11)	(1·20)			36·0		
CO_2	33·0	(1·16)	(1·28)					
C_2H_2								
C_2H_4								
CH_4	29·78	(1·05)						
C_3H_8								
SO_2								

Table 6 (*cont.*)

	Gerbes (1937)		Valentine and Curran (1952)		Jesse and Sadauskis (1955)	Weiss and Bernstein (1955, 1956)		Hardwick (unpubl.)	Dorfman (1954)
Gas	w (ev)	w_{gas}/w_A	w (ev)	w_{gas}/w_A	w_{gas}/w_A	w (ev)	w_{gas}/w_A	w_{gas}/w_A	w_{gas}/w_A
He			32·5	(1·20)	(1·61)	40·3	(1·57)	(1·16)	(1·28)
Ne	29·1	(1·07)			(1·39)	35·3	(1·48)		
A	27·3	(1·00)	27·0	(1·00)	(1·00)	25·8	(1·00)	(1·00)	(1·00)
Kr					(0·915)	24·7	(0·96)		
Xe					(0·83)	22·0	(0·82)		
H_2			38·0	(1·41)	(1·34)	36·3	(1·40)	(1·41)	(1·45)
Air	33·4	(1·22)	35·0	(1·29)	(1·29)	33·9	(1·31)	(1·28)	
N_2	34·6	(1·27)	35·8	(1·32)	(1·31)	34·6	(1·34)	(1·33)	
O_2	30·5	(1·12)	32·2	(1·19)	(1·17)	31·2	(1·21)	(1·18)	(1·18)
CO_2	32·2	(1·18)			(1·25)	32·6	(1·26)		
C_2H_2					(0·99)	26·4	(1·02)		
C_2H_4					(0·94)				
CH_4			29·5	(1·09)	(1·03)	26·8	(1·04)	(1·05)	
C_3H_8	23·6	(0·86)							
SO_2	30·4	(1·11)							

Relative values in the form w_{gas}/w_{air} are in brackets.

§ 4. Heavy Particles

4.1. *Alpha Particles*

Absolute values of w have been determined for the ionization of a wide range of gases by α-particles. Here there have not been as many fundamentally different methods employed as in the case of ionization by electrons. The most important distinction lies in the methods used to collect ions. One group of experiments is based on slow ion collection where the positive ions play an important part and the other group makes use of fast gridded chambers in which only the electronic component of the ionization is collected.

4.1.1. *Slow chambers.*

The early work covered in this section has been reviewed by Gray (1944). A typical experiment in this group involved the use of the α-particles of RaC′ which have an energy of 7·683 MeV. The strength of the source was measured in terms of the γ-ray activity of Ra (B + C) with which the RaC′ was in equilibrium. The number of α-particles emitted per second per gramme of radium is known to be $3{\cdot}7 \times 10^{10}$ and determination of the total ionization caused by the RaC′ α-particles completes the data necessary for calculation of the value of w. Gray quotes the value of $w_{air} = 35{\cdot}1$ ev for RaC′ α-particles from the work of Fonovits-Smereker (1922), Curie and Joliot (1928), Bracelin (unpublished) and Jentschke (1935). Values of w for other gases may be found relative to air, and on the basis of results of Gurney (1925) and Taylor (1911, 1913) Gray concludes that the best value for hydrogen is 36·0 ev for Po α-particles of energy 5·3 MeV. A value for helium was measured by Gurney (1925) and also by Naidu (1936) and the best compromise was obtained with $w_{He} = 31{\cdot}0$ ev.

Relative values were measured by Schmieder (1939) and Stetter (1943), who determined saturated ionization currents in various gases relative to air as a standard. These are listed in table 7. Schmieder converted the ionization chamber measurements to absolute values by using a value of $5{\cdot}22 \times 10^6$ ev for the energy of a Po α-particle together with a value of $1{\cdot}5 \times 10^5$ for the number of ions created in air by a single particle. This latter figure is due to Meyer and Schweidler (1927). Absolute values were calculated from Stetter's work by Binks (1954) and these are also shown in table 7.

Table 7

	Schmieder		Stetter	
Gas	w (ev)	w_{gas}/w_{air}	w (ev)	w_{gas}/w_{air}
Air	34·8	1·000	35·7	1·000
N_2	36·1	0·965	37·2	0·96
O_2	32·4	1·075		
NO	30·05	1·135		
NO_2	33·0	1·052		
N_2O	32·4	1·075		
CO	34·0	1·025		
CO_2	33·9	1·028		
C_2H_4	27·0	1·290		
Ne	27·2	1·28		1·20
A	24·37	1·43	28·3	1·26
Kr	22·43	1·55		1·36
Xe				1·51
H_2			36·1	0·99

Dick, Falk-Vairant and Rossel (1947) determined w for N_2, O_2 and CH_4 relative to air; their absolute values are much lower than those obtained by other workers due to the fact that the value they obtained for w_{air} was 32·1 ev.

Very much better agreement is shown between the values obtained by several different experimenters in more recent years. Reference to table 8 shows that for all gases, with the one exception He, the absolute values of w agree to within about ± 0·5 ev. The methods of measurement differ in detail but are all based

Table 8

Gas	Jesse and Sadauskis (1953)		Sharpe (1952)		Valentine and Curran (1952)		Huber and co-workers† (1953, 1955)		Bortner and Hurst (1954)	
	w (ev)	w_{gas}/w_A	w (ev)	w_{gas}/w_A	w (ev)	w_{gas}/w_A	w (ev)	w_{gas}/w_A	w (ev)	w_{gas}/w_A
A	26·4	(1·00)	26·3	(1·00)	25·9	(1·00)	26·25	(1·00)	26·4	(1·00)
He	42·7	(1·617)			31·7	(1·22)	29·6	(1·13)	46·0	(1·74)
Ne	36·8	(1·394)								
Kr	24·1	(0·913)								
Xe	21·9	(0·830)								
H_2	36·3	(1·375)			37·0	(1·43)	35·96	(1·36)	37·0	(1·40)
CO_2	34·5	(1·307)	34·2	(1·300)			34·3	(1·31)	34·3	(1·30)
Air	35·5	(1·345)	35·6	(1·354)	35·2	(1·36)	34·95	(1·33)	35·0	(1·33)
O_2	32·5	(1·231)	32·9	(1·251)	32·2	(1·24)	32·17	(1·275)	32·2	(1·22)
N_2	36·6	(1·386)	36·4	(1·384)	36·0	(1·39)	36·5	(1·39)	36·3	(1·37)
CH_4	29·2	(1·106)	29·1	(1·106)	29·0	(1·12)	29·0	(1·105)	29·4	(1·11)
C_2H_2	27·5	(1·042)								
C_2H_6	26·6	(1·008)								
C_2H_4	28·4	(1·061)							28·0	(1·06)
BF_3							35·3	(1·36)	36·0	(1·36)
SF_6									35·7	(1·35)
Freon 12									29·5	(1·12)
H_2S							23·4	(0·89)		
NH_3							30·5	(1·16)		
C_4H_{10}							23·0	(0·875)		
SO_2							32·5	(1·24)		
CCl_4							25·9	(0·99)		
C_2H_5OH							32·6	(1·24)		

† Biber, Huber and Müller (1955) and Haeberli, Huber and Baldinger (1953).

on 'slow' ionization chambers, and for these conditions there are no major discrepancies such as appear in the electron results. The Argonne measurements, reported by Jesse and Sadauskis (1953), were carried out on single α-particles.

The very marked discrepancies shown in the case of helium for which the absolute values range from about 30 ev to 43 ev have been shown to be due to impurities and this phenomenon is discussed in more detail in § 5.1.

4.1.2. *Fast chambers.*

It is possible to measure the ionization produced by α-particles with a gridded chamber. In this case only the 'fast' electronic component of the ionization is collected. This method was introduced by Cranshaw and Harvey (1948) and Hanna (1950), but these workers were mainly concerned with the variation of w_A with energy and their results are discussed in § 4.1.3. Absolute values of w were obtained, by means of gridded chambers, by Herwig and Miller (1954 a) and Bertolini, Bettoni and Bisi (1952), as shown in table 9.

Table 9

	Herwig and Miller (1954)		Bertolini *et al.* (1952)	
Gas	w (ev)	w_{gas}/w_A	w (ev)	w_{gas}/w_A
A	27·2	(1·000)	28·9	(1·00)
He	44·2	(1·62)		
K	25·7	(0·945)		
Ne	38·6	(1·41)		
N_2	38·8	(1·43)	37·9	(1·31)

Herwig and Miller took similar precautions to those of Jesse and Sadauskis (1953) concerning the purity of the gases and they obtained high values for helium and neon. The relative values obtained are in general similar but the absolute values are about 6% higher in the case of the fast collection. A qualitative explanation of this phenomenon can be given in terms of energy which is temporarily stored in excited states and therefore lost in the case of fast electron collection, but at present there is no satisfactory quantitative explanation which can explain a difference of this magnitude.

4.1.3. *Variation with energy.*

Early work on the variation of w for air with energy of α-particles has been reviewed by Gray (1944). w can be considered as the ratio of the rate of loss of energy by the particle to the ionization produced per unit path at a given point in the range, i.e. $w = (dE/dx)/(dI/dx)$. A comparison of the instantaneous values of dE/dx and dI/dx over the range of the α-particles then leads to a variation of w with energy. Difficulties are encountered at low energies because the α-particle beam becomes inhomogeneous in energy and the values for dE/dx become somewhat unreliable. After consideration of several experiments of this type, Gray came to the conclusion that w_{air} varied with α-particle energy and that this variation could be represented by a formula similar to that proposed by Gerbes for electrons but with modified values of the constants. It should be noted however

that most of the actual experimental results quoted by Gray were for the energy range above 1 MeV where only a very small variation of the order of 1–2% would be expected from a Gerbes type formula.

The variation of w for argon with energy was investigated by Cranshaw and Harvey (1948) using a gridded chamber. Their experimental results showed that the pulse size varied linearly with energy in the range investigated, i.e. 5 to 9 MeV. Extrapolation of the ionization–energy line gave an energy of 85 kev for zero ionization. This intercept on the energy axis was taken by the authors to be an indication of a variation of w at low energies. Cranshaw and Harvey then fitted a Gerbes type formula to their experimental results and decided that these were best summarized by the formula $w = 27{\cdot}5 + 1{\cdot}9/E^{1/2}$.

A very useful reaction for obtaining information concerning w values for low energy α-particles and also ^{7}Li ions is ^{10}B(n, α)^{7}Li. The reaction energy is known to be $2{\cdot}793 \pm 0{\cdot}003$ MeV from precisely measured Q values for ^{10}B(p, α)^{7}Be and ^{7}Li(p, n)^{7}Be. When the lithium nucleus is left in an excited state the energy released is less than the above value by 479 ± 1 kev, this being the energy of the γ-ray emitted by the excited nucleus when it returns to the ground state. The energy of each of the particles is known; ionization current measurements then enable w values to be found.

Hanna (1950) used a gridded ionization chamber with a filling of argon + 2% BF_3 to study this reaction. The results, shown in table 10, column 4, are in good agreement with the values predicted by the Cranshaw–Harvey formula (column 3). A similar experiment with a chamber filling of A + 5% CO_2 was carried out by Rhodes, Franzen and Stephens (1952). As can be seen from column 5, the measurements for α-particles agree well with those in columns 3 and 4 but the Li values indicate a slightly greater variation in w than that given by the formula.

Table 10

Particle	Calculated energy (MeV)	w relative to fast α-particle value: C.–H. formula	Hanna	Rhodes *et al.*
α	$1{\cdot}778 \pm 0{\cdot}002$	1·021		$1{\cdot}021 \pm 0{\cdot}010$
α*	$1{\cdot}473 \pm 0{\cdot}002$	1·026	$1{\cdot}027 \pm 0{\cdot}009$	$1{\cdot}020 \pm 0{\cdot}007$
Li	$1{\cdot}015 \pm 0{\cdot}002$	1·059		$1{\cdot}091 \pm 0{\cdot}015$
Li*	$0{\cdot}041 \pm 0{\cdot}002$	1·068	$1{\cdot}080 \pm 0{\cdot}011$	$1{\cdot}094 \pm 0{\cdot}009$

Strictly, the figures quoted from Hanna are values relative to the 5·159 MeV α-particle from ^{239}Pu, while those from Rhodes *et al.* are relative to the 5·298 MeV Po α-particle. This makes no difference to the general conclusion that the value of w in argon increases for decreasing α-particle energy.

The results from experiments of this type can be displayed in a different form. The energy ratio E_α/E_{Li} can be precisely calculated from atomic masses and is found to be 1·7529; this ratio can also be determined experimentally by ionization measurements, and the results of several workers are given in table 11.

These ratios, which are consistent with one another, can only be explained by assuming that the value of w in argon depends on the velocity or charge of the ionizing particle. This result is reinforced by work reported very recently by Curtis and Miller (1956). The same reaction, ^{10}B(n, α)^{7}Li was studied and it

was shown that the ratio $w_{Li}/w_\alpha > 1$. Ionization defects of 51 ± 3 and 61 ± 3 kev were obtained for the Li recoils in $A + N_2$ and $A + CO_2$ mixtures respectively. (Ionization defects are discussed in detail in § 4.3.)

Table 11

Name	E_α/E_{Li}
Hanna (1950)	$1{\cdot}841 \pm 0{\cdot}025$
Stebler, Huber and Bichsel (1949)	$1{\cdot}89 \pm 0{\cdot}02$
Rhodes *et al.* (1952)	$1{\cdot}878 \pm 0{\cdot}014$
De Juren and Rosenwasser (1953)	In agreement with above

The conclusions stated above concerning the variation of w_A with α-particle energy are in marked contrast to the findings of Jesse, Forstat and Sadauskis (1950). These workers used an interesting method in which the ionization of a single α-particle was recorded, after amplification, by a vibrating reed electrometer. No variation in w_A with energy was found. The energy range 5 to 9 MeV was investigated directly, using α-particles from naturally radioactive elements and the region 1–5 MeV by an indirect method involving the use of α-particles from nuclear reactions. Earlier a comparison of the relative ionization of Po and Sm α-particles in argon by Jesse and Sadauskis (1949 a, b) had shown that the ratio of ionization $(Sm/Po)_{argon}$ is about 5% higher than the ratio $(Sm/Po)_{air}$. This result, coupled with the constancy of w in argon obtained by these workers, led them to the conclusion that w_{air} is not constant. This deduction is indirectly confirmed by later work reported by Jesse and Sadauskis (1955). The experiment has already been discussed in § 3.2.1, where it was noted that relative values of w for ionization by electrons was found for a large series of gases. These gave no indication of a variation with energy when the ionization was caused by β-particles for ^3H or ^{63}Ni. The value for β-particles, w_β, was compared with values for α-particles of different energies (5 MeV or less). Jesse and Sadauskis showed that the ratio w_α/w_β was constant for hydrogen and the noble gases but was not constant for air and the other gases investigated. They therefore concluded that w_α was independent of α-particle energy in the case of the noble gases and hydrogen. For air and the other gases w_α appears to increase with decreasing energy of the α-particles. A tendency is noted for w_α to reach a constant value for these gases at high α-particle energies.

It is interesting to note that, as in the case of the absolute values, there is a distinct difference in the results obtained with fast and slow methods of ion collection.

4.2. *Protons*

The experimental work on w for protons is much less complete than for α-particles.

An early measurement for very low energy protons is due to Gerthsen (1930). He used an apparatus very similar to that employed by Eisl (1929) for electrons. The protons were produced in an evacuated canal tube and introduced into the ionization chamber through a thin 'Cellophane' window. The energy of the beam was determined by magnetic deflection. Gerthsen reported that for air an absolute value of $w = 36$ ev was obtained and no significant variation was found in the

energy range 27 to 45 kev. A similar value was found for hydrogen. Jentschke (1940) employed a linear amplifier to investigate the ionization–range relation for individual protons in nitrogen, and Gray (1944) deduced a value of $w_{air} = 35{\cdot}1$ ev for protons of initial energy of 7·56 Mev.

The relative ionization produced by low energy protons, deuterons and α-particles was studied by Tunnicliffe and Ward (1952). The particles were provided by neutron recoils and the ionization measurements were made with an all-glass Maze-type proportional counter. Their results show that the value of w for protons of energy 200–550 kev in argon and argon–methane mixtures is constant to within 1% in the energy range investigated. A further conclusion from this work is that the value of w for protons is not the same as that for α-particles of comparable energy. In fact, the total ionization in argon produced by a proton is about 8% greater than an α-particle of equal energy. This might indicate that the value of w is dependent on the nature of the ionizing particle although it could reflect the difference in values obtained with particles of equal energy but different velocities.

Very high energy protons of 340 Mev were used by Bakker and Segrè (1951) to measure the values of w relative to w_A for various gases. In this work the proton beam was made to cross two identical ionization chambers one of which was filled with argon at atmospheric pressure and served as a monitor. The other was successively filled with the gases under investigation and w was calculated in the following way :

$$w = -\frac{dE}{dN} = -\frac{dE/dx}{dN/dx}.$$

Values of the term $-dE/dx$ were obtained for each gas from theoretical tables and since dN/dx was proportional to the ionization measured in the chamber, values of w relative to argon could be calculated. To convert these relative values to absolute values a measurement (by Chamberlain, Segrè and Wiegand, unpublished) of the number of ion pairs produced by one 340 Mev proton crossing 1 cm of argon at atmospheric pressure was used to give dN/dx absolutely for argon. This, together with the theoretical value of $-dE/dx$, gave $w_A = 25{\cdot}5$ ev. The full series of results are shown in table 12.

Table 12

Gas	w (ev)	w_{gas}/w_A
A	25·5	1·00
He	28·8	1·17
H_2	35·2	1·38
N_2	33·4	1·31
Air	33·2	1·30
O_2	31·2	1·23

This set of results is the most comprehensive so far published for protons. The relative values are in general more similar to those shown in table 3 for electrons than to those in table 8 for α-particles. There are however too few data available for protons to deduce that the values of w are significantly different from those of α-particles. The experimental evidence is not sufficiently detailed to allow any comment to be made on a possible variation of w with proton energy.

4.3. *Heavy Recoil Particles and Fission Fragments*

The ionization of a heavy particle is very critically dependent on its velocity. The velocity v is conveniently discussed in terms of $v_0 = e^2/\hbar$. At high velocities, $v \gg v_0$, the heavy particle loses energy primarily to electrons and w is close to the value found for α-particles. This fast particle value for w may be designated w_α. At low velocities, $v \sim v_0$, energy loss to recoiling atoms predominates because of the neutralization of the particle charge through the capture of electrons. If these secondary particles have a low overall efficiency for the production of ion pairs the value for w for the primary particle will be greatly increased. The theoretical aspects of this problem have been investigated by Knipp and Ling (1950, 1951) using the concept of ionization defect. The ionization defect ΔE is defined as $\Delta E = E - w_\alpha I$, where E is the energy of the primary particle and I the amount of ionization it produces. ΔE is then a measure of the degree to which the total ionization, when converted to energy units by the factor w_α, fails to give a true measure of the particle energy. Knipp and Ling (1951) developed expressions for the numerical estimation of ΔE, but due to lack of information concerning the ratio of the stopping cross sections for velocities below v_0, these estimates are considered by the authors to be crude.

For the ionization of argon gas by argon particles of very high velocity they obtain ΔE of the order of 780 kev. In the case of ionization of argon gas by fission fragments the values of ΔE are

$$\Delta E_L \sim 2{\cdot}5 \text{ Mev for a light fragment of energy 98 Mev,}$$

$$\Delta E_H \sim 4{\cdot}2 \text{ Mev for a heavy fragment of energy 67 Mev.}$$

There is not a great deal of experimental evidence concerning ionization by heavy particles. Some of the earlier results were presented in the form w_R/w_α where w_R refers to a recoil particle. Wertenstein (1913) (quoted by Madsen (1945)) obtained a value between 2 and 8 for this ratio when applied to ionization of air by RaC recoil particles. Gerthsen and Grimm (1943) used ThC′ recoil particles in ionization chambers and found the ratio to be about unity for air. A more comprehensive investigation was undertaken by Madsen (1945). The ionization of single recoil atoms from Po, ThC and ThC′ was studied using a proportional counter and coincidence arrangement. The proportional counter was filled with argon + 5% air and the relative amount of ionization produced by the recoil atoms was compared with that produced by α-particles of known energy. For this gas mixture the ratio w_R/w_α was found to be 4·4, 3·8 and 3·4 for Po, ThC and ThC′ respectively. Madsen calculated the value of $w_R = dE/dI$ for heavy particles in the energy range 100–170 kev. This was found to be 67 ev relative to $w_\alpha = 25$ ev.

Information suitable for quantitative comparison with the calculations of Knipp and Ling is difficult to obtain although disagreements between ionization and calorimetric measurements on the energy of fission fragments indicated the existence of an ionization defect. An absolute determination of ΔE requires a measurement of the energy of the fragment, E_f, by a method independent of ionization. In a comparison experiment Leachman (1951) found the ionization defect for light fragments to be approximately 3·7% less than that for heavy

fragments. The same author (Leachman 1952) made an absolute measurement of ionization defects using a method in which velocity distributions obtained by a time-of-flight technique were compared with those from ionization chamber data. The results for an A+3%CO_2 mixture are shown in table 13 together with w values for the light and heavy fragments relative to the α-particles emitted by the uranium and plutonium isotopes.

Table 13

	^{233}U	^{234}U	^{239}Pu	Knipp and Ling (1951)
ΔE_L (MeV)	6·1	5·7	5·2	2·5
ΔE_H (MeV)	7·3	6·5	6·4	4·2
w_L/w_α	1·06	1·06	1·05	—
w_H/w_α	1·13	1·11	1·10	—

Although the experimental values of ΔE are higher than those calculated by Knipp and Ling (1951) they are successful in reconciling the results of ionization chamber experiments and the data for calorimetric and mass spectrographic measurements. The close agreement between many ionization chamber measurements by different investigators suggests that ionization defects are real and inherent in the ionization process rather than merely experimental errors due to recombination and similar effects.

This study was extended to other gases by Herwig and Miller (1954 b), who used a gridded chamber to measure the ionization produced by fission fragments relative to that produced by the α-particles of uranium. No attempt was made to measure E_f absolutely; instead the difference in ΔE for two gases was found. E_f can be expressed as before, $E_f = w_\alpha I_f + \Delta E_f$, and if two gases are considered it can be eliminated, leaving

$$(\Delta E_f)_1 - (\Delta E_f)_2 = (w_\alpha I_f)_2 - (w_\alpha I_f)_1.$$

The quantities on the right-hand side of this equation are either known or can be measured. The method consisted of measuring the ratio of the most probable ionizations for fission fragments and α-particles in various gases. The results are shown in table 14. Since A+3%CO_2 has been used in previous work (including that of Leachman described above) the ionization defects for the other gases are calculated relative to ΔE for this mixture.

Table 14

Gas	$(w_\alpha I_f)_\alpha$ (MeV)		$(w_\alpha I_f)_\alpha - (w_\alpha I_f)_{A+3\%\,CO_2}$ (MeV)	
	light	heavy	light	heavy
He	92·9 ± 0·5	61·8 ± 0·7	3·4 ± 0·7	5·4 ± 1·0
Ne	92·0 ± 0·5	60·0 ± 0·7	2·5 ± 0·7	3·6 ± 1·0
A	92·5 ± 0·5	60·9 ± 0·7	3·0 ± 0·7	4·5 ± 1·0
Kr	89·6 ± 0·7	56·0 ± 1·0	0·1 ± 0·9	−0·4 ± 1·2
N_2	88·9 ± 0·5	58·0 ± 0·7	−0·6 ± 0·7	1·6 ± 1·0
A+3% CO_2	89·5 ± 0·5	56·4 ± 0·7	0	0
A+5% N_2	91·5 ± 0·5	59·9 ± 0·7	2·0 ± 0·7	3·5 ± 1·0

It can be seen that ΔE for argon is much the same as that in helium, a result which was unexpected on theoretical grounds. Possible explanations of this are

discussed in detail by Herwig and Miller. Another surprising effect is the large difference caused by the addition of 3% CO_2 to argon. This is difficult to understand particularly since the addition of 5% of N_2, itself a gas with a large defect, makes only a small difference when added to argon. Measurements being carried out by Miller and his co-workers on ionization defects of slow α-particles in various gases may help to provide explanations to some of these problems.

§ 5. Mixtures of Gases

5.1. *Low Concentrations of One Gas in Another*

Early experiments designed to determine the absolute value of w for helium yielded results of the order of 30 ev and similar values were obtained in later work, such as that reported by Valentine and Curran (1952) in which small amounts of a quenching gas were deliberately added to the helium. The first important departure from this value was reported by Jesse and Sadauskis (1952, 1953), who showed that the value obtained for w_{He} was very critically dependent on its purity. Helium which was impure due to a concentration of argon of only about 0·13% yielded a value of w_{He} of the order of 30 ev and this steadily increased as the amount of contaminant decreased. For helium of the highest purity a maximum value of 42·7 ev was obtained. The increased ionization when an impurity was present was attributed to the fruitful destruction of the metastable states in the helium according to the equation $He^* + A \rightarrow He + A^+ + e^-$. This reaction can take place since the ionization potential of argon is less than the excitation energy of the metastable state in helium. Krypton and carbon dioxide as contaminants in helium yielded similar results. Neon was also found to have similar properties to helium, the values of w increasing from 26·1 ev for neon with 0·12% argon contaminant to 36·8 ev for very pure neon. Similar results have since been reported by other workers, notably Bortner and Hurst (1954).

As shown in § 2, the most recent theoretical value for w_{He} obtained by Erskine (1954 a) is in good agreement with the experimental value for very pure helium. Erskine (1954 b) made a further calculation to test the validity of the hypothesis that the low value obtained in the presence of minute amounts of contaminant was due to the de-activation of long lived metastable states in helium atoms. He was able to show that the experimental results could be explained by a process described by the equation $He^* + A \rightarrow He + A^+ + e^-$ only by assuming that *all* excited helium atoms lead eventually to ion pairs. The alternative assumption that only metastable states were involved did not lead to a sufficient reduction in the value of w.

This conclusion has indirect support from the work of Sharpe (1952), who obtained a similar increase in the total ionization produced in argon by Pu α-particles when very small amounts of carbon dioxide were added. In this case the increase cannot be attributed, even partially, to the de-excitation of metastable states since the ionization potential of carbon dioxide is greater than the excitation energy of the metastable state in argon. Further confirmation of this result was reported by Bertolini, Bettoni and Bisi (1954), who investigated the ionization of gas mixtures by Po α-particles using a gridded ionization chamber. The ionization of argon with small traces of H_2, N_2, CH_4, C_2H_5OH or C_6H_6 were particularly

studied. No increase in the total ionization for argon was obtained for small admixtures of H_2 and N_2 but an increase was reported in the case of the other three gases. Bertolini *et al.* ascribed the results as due to the fact that the ionization potentials of A, N_2 or H_2 are greater than the excitation energies of any levels in either A, N_2 or H_2. Thus no new ions could be created by de-excitation processes. In the case of the other gases the following reaction was assumed to have taken place, $A^* + S \rightarrow S^+ + A + e^-$, where A^* is an excited argon atom and S an atom of either CH_4, C_2H_5OH or C_6H_6 which all have ionization potentials less than the excitation energy of argon. It was found that the lower the ionization potential of the foreign gas the higher the maximum ionization.

This study was extended by Melton, Hurst and Bortner (1954) using apparatus designed by Bortner and Hurst (1954) in which the ionization of α-particles from ^{239}Pu was recorded with a large parallel plate ionization chamber. The rate of ionization was measured by the voltage change on an accurately calibrated capacitor and the voltage change was determined by means of a vibrating reed electrometer used as a null indicator. The ^{239}Pu source was electroplated on platinum and used without collimation. The total ionization in argon in the presence of a small amount of a foreign gas was measured for a large number of gases. In agreement with previous work it was found that any gas with an ionization potential of less than 15 ev (the largest amount of excitation energy available from argon) would produce a decrease in the value of w when added to argon. In conflict with the results of Bertolini *et al.* (1954) little correlation was found between the ionization potential of the gas and the amount of decrease in w, except that in general the effect was greater if the ionization potential was less than the 11·5 volt metastable state in argon. Using a measurement due to Biondi (1952), Melton *et al.* showed that de-excitation of an excited state would explain the observed increases in ionization if the lifetime of the state was only of the order of a few microseconds.

5.2. *Complete Mixtures of Two Gases*

The first report of a study of the ionization of a gas mixture from 100% of one gas to 100% of the other gases is due to Gerbes (1937). He investigated the ionization of a propane–nitrogen mixture by electrons of energy 10–40 kev. No further work on this subject was done until the ionization of gas mixtures by α-particles was discussed by Huber, Baldinger and Haeberli (1950). These authors derived an expression for the variation of w_{mixture} with the partial pressures of the component gases,

$$\frac{1}{w_m} = \frac{s_1 p_1}{s_1 p_1 + s_2 p_2}\left(\frac{1}{w_1} - \frac{1}{w_2}\right) + \frac{1}{w_2} \qquad \ldots\ldots(1)$$

where s_1, s_2 and p_1, p_2 are the stopping powers and partial pressures of the component gases. This expression is obtained by considering that the gases act independently of one another. In particular the ionizing effect of the delta electrons and of the light quanta of one gas with respect to the other is completely neglected. Equation (1) shows that, to a first approximation, $1/w_m$ should be linearly dependent on $Z = s_1 p_1/(s_1 p_1 + s_2 p_2)$. The variation of $1/w_m$ with Z was plotted experimentally by Haeberli, Huber and Baldinger (1953) for several gas

mixtures and the results were subsequently extended by other authors. It was found that the linear relationship between $1/w_m$ and Z held good for only a few gas mixtures. To account for the differences between theory and experiment, Haeberli *et al.* (1953) derived a modified expression for $1/w_m$,

$$\frac{1}{w_m} = \left(\frac{1}{w_1} - \frac{1}{w_2}\right) Z + \frac{1}{w_2} + \Delta. \qquad \ldots\ldots(2)$$

Δ is a complicated expression which takes into account δ-rays produced in gas (1) ionizing in gas (2) and vice versa. It is defined in terms of the mean stopping powers for δ-rays in the two gases and similar factors which are not explicitly determinable. By suitable choice of constants Haeberli *et al.* were able to fit their experimental results to curves derived for Δ.

An alternative approach has been given by Bortner and Hurst (1954). If it is assumed that the 'effective' stopping power for insertion into equation (1) is not the actual stopping power but differs from s_1 by Δs, then an equation exactly similar in form to that obtained by Haeberli *et al.* for Δ can be derived. An empirical expression for gas mixtures which accounts for the difference in w value due to δ-rays and also the differences in effective stopping powers is suggested by Bortner and Hurst as a result of this calculation. It is

$$\frac{1}{w_m} = \left(\frac{1}{w_1} - \frac{1}{w_2}\right) Z + \frac{1}{w_2} \qquad \ldots\ldots(3)$$

where

$$Z = \frac{p_1}{p_1 + a p_2}.$$

The constant a is determined empirically for every mixture and is equal to s_2/s_1 in cases where equation (1) holds. Curves given by Bortner and Hurst show good agreement with this relation when the correct value of a is used.

It is possible to deduce a different expression for w_m without making the assumption that the gases behave independently of one another. This has been done by Valentine (1953) by considering the total number of ions in the gas mixture as the sum of the ions produced by primary ionizations in each gas component and by secondary δ-rays, i.e. $\Delta n = \Delta n_1 + \Delta n_2 + \Delta n_s$. The expression for w_m derived in this way can be written in the form

$$\frac{1}{w_m} = \frac{1}{w_e} + \frac{1}{w_{p2}}\left(1 - \frac{\overline{I}_2}{w_e}\right) + Z\left[\frac{1}{w_{p1}} - \frac{1}{w_{p2}} - \frac{1}{w_e}\left(\frac{\overline{I}_1}{w_{p1}} - \frac{\overline{I}_2}{w_{p2}}\right)\right] \qquad \ldots\ldots(4)$$

where $w_e = w$ value for δ-rays in the gas mixture, $w_{p1}, w_{p2} = w$ values for primary α-particle ionization in the two gas components, $\overline{I}_1, \overline{I}_2$ = mean ionization potentials of two gas components, $Z = s_1 p_1/(s_1 p_1 + s_2 p_2)$ as before.

Experimental information interpreted by means of equation (4) has been published by Valentine (1952, 1953). This consists of curves in which $1/w_m$ is plotted against Z for A–H_2, A–N_2 and A–He mixtures. These curves are available for slow electrons from ^{37}A and for Po α-particles. Since the energy of the ^{37}A electrons is of the same order of magnitude as the secondary δ-rays produced by the α-particles the electron curves may be regarded as secondary ionization curves and the α curves as total (primary + secondary) curves. In equation (4) w_e is the

dominant term and thus to a first approximation $w_m = w_e$ for a given Z. Thus when $1/w_m$ is plotted against Z the α-particle and electron curves should be nearly coincident. The remaining terms give the amount of separation of the curves. Were these terms constant the two curves would be parallel; since they contain w_e which is dependent on Z the separation will vary with Z. Furthermore the correction term is always positive and the α-particle curve would be expected to lie above the electron curve in every case. These conclusions are justified by the published curves.

Information about a pure gas can be obtained by substituting $Z = 0$ or 1 in equation (4) which then reduces to

$$\frac{1}{w_g} = \frac{1}{w_e} + \frac{1}{w_{p1}}\left(1 - \frac{\bar{I}_1}{w_e}\right). \qquad \ldots\ldots(5)$$

For a pure gas w_p is expected to be about 100 ev (Mott and Massey 1949). For pure argon, substituting this value together with the experimentally determined value of w for electrons and α-particles a value $\bar{I}_A \sim 23$ ev is obtained. This is greater than the minimum ionization potential but is similar to the value $\bar{I}_A \sim 25$ ev deduced by Townsend (1947) for an average ionization potential for slow electrons in air.

The expressions derived by different authors are useful as methods of summarizing experimental results. They do not, however, contribute much to the understanding of the ionization of gas mixtures.

References

Anslow, G., 1925, *Phys. Rev.*, **25**, 484.

Bagge, E., 1937, *Ann. Phys., Lpz.*, **30**, 72.

Bakker, C. J., and Segrè, E., 1951, *Phys. Rev.*, **81**, 489.

Bertolini, G., Bettoni, M., and Bisi, A., 1952, *Nuovo Cim.*, **9**, 1004; 1954, *Ibid.*, **11**, 458.

Bertolini, G., Bisi, A., and Zappa, L., 1953, *Nuovo Cim.*, **10**, 1424.

Bethe, H., 1930, *Ann. Phys., Lpz.*, **5**, 325.

Biber, C., Huber, P., and Müller, A., 1955, *Helv. Phys. Acta*, **28**, 503.

Binks, W., 1936, *Rep. Progr. Phys.*, **3**, 347 (London: Physical Society); 1954, *Acta Radiol.*, Suppl., **117**, 85.

Biondi, M. A., 1952, *Phys. Rev.*, **88**, 660.

Bohr, N., 1948, *K. Danske Vidensk. Selsk., Mat.-fys. Medd.*, **18**, No. 8.

Bortner, T. E., and Hurst, G. S., 1954, *Phys. Rev.*, **93**, 1236.

Breunig, E., 1936, *Ann. Phys., Lpz.*, **25**, 467.

Broyles, C. D., Thomas, D. A., and Haynes, S. K., 1953, *Phys. Rev.*, **89**, 715.

Buchmann, E., 1928, *Ann. Phys., Lpz.*, **87**, 509.

Cockroft, A. L., and Curran, S. C., 1951, *Rev. Sci. Instrum.*, **22**, 37.

Compton, A. H., and Allison, S. K., 1936, *X-rays in Theory and Experiment* (New York: Van Nostrand).

Cranshaw, T. E., and Harvey, J. A., 1948, *Canad. J. Res.* A, **26**, 243.

Crowther, J. A., and Orton, L. H. H., 1930, *Phil. Mag.*, **10**, 329; 1932, *Ibid.*, **13**, 505.

Curie, I., and Joliot, F., 1928, *C.R. Acad. Sci., Paris*, **187**, 43.

Curran, S. C., Angus, J., and Cockroft, A. L., 1949, *Phil. Mag.*, **40**, 36.

Curran, S. C., Cockroft, A. L., and Insch, G. M., 1950, *Phil. Mag.*, **41**, 517.

Curtis, D. W., and Miller, G. H., 1956, *Nuclear Sci. Abstr.*, **10**, No. 15.

De Juren, J., and Rosenwasser, H., 1953, *Phys. Rev.*, **92**, 544.

Dick, L., Falk-Vairant, P., and Rossel, J., 1947, *Helv. Phys. Acta*, **20**, 357.

Dorfman, L. M., 1954, *Phys. Rev.*, **95**, 393.

EISL, A., 1929, *Ann. Phys., Lpz.*, **3**, 277.
EMERY, E. W., 1956, *Brit. J. Radiol.*, **29**, 370.
ERSKINE, G. A., 1954 a, *Proc. Roy. Soc.* A, **224**, 362 ; 1954 b, *Proc. Phys. Soc.* A, **67**, 640.
FAILLA, G., 1956, *Proceedings of International Conference on Peaceful Uses of Atomic Energy*, Vol. 14, 239 (New York : United Nations).
FAILLA, G., and ROSSI, H. H., 1952, *Annual Report on Research Project*, Report NYO 4008 (U.S. Atomic Energy Commission).
FANO, U., 1946, *Phys. Rev.*, **70**, 44.
FONOVITS-SMEREKER, H., 1922, *S.B. Akad. Wiss. Wien*, **131**, IIa, 355.
FREUND, L., 1935, *Ann. Phys., Lpz.*, **22**, 748.
FROST, R. H., and NIELSEN, C. E., 1953, *Phys. Rev.*, **91**, 864.
FULBRIGHT, H., 1955, *Beta- and Gamma-ray Spectroscopy*, ed. by K. Siegbahn (Amsterdam : North-Holland).
GAERTNER, O., 1929 a, *Ann. Phys., Lpz.*, **2**, 94 ; 1929 b, *Ibid.*, **3**, 329 ; 1931 a, *Ibid.*, **10**, 825 ; 1931 b, *Ibid.*, **11**, 648 ; 1933, *Ibid.*, **16**, 613 ; 1934/5, *Ibid.*, **21**, 564 ; 1935, *Ibid.*, **23**, 255.
GERBES, W., 1935, *Ann. Phys., Lpz.*, **23**, 684 ; 1937, *Ibid.*, **30**, 169.
GERTHSEN, C., 1930, *Ann. Phys., Lpz.*, **5**, 657.
GERTHSEN, C., and GRIMM, Z., 1943, *Z. Phys.*, **120**, 476.
GRAY, L. H., 1936, *Proc. Roy. Soc.* A, **156**, 578 ; 1944, *Proc. Camb. Phil. Soc.*, **40**, 72.
GROSS, W., 1954, *Annual Report on Research Project*, Report NYO 4582 (U.S. Atomic Energy Commission).
GURNEY, R. W., 1925, *Proc. Roy. Soc.* A, **107**, 332.
HAEBERLI, W., HUBER, P., and BALDINGER, E., 1953, *Helv. Phys. Acta*, **26**, 145.
HANNA, G. C., 1950, *Phys. Rev.*, **80**, 530.
HARRISON, G. R., CRAWFORD, R. C., and HOPKINS, J. I., 1955, *Phys. Rev.*, **100**, 841.
HEINTZE, J., 1955, *Z. Phys.*, **143**, 153.
HERWIG, L. O., and MILLER, G. H., 1954 a, *Phys. Rev.*, **94**, 1183 ; 1954 b, *Ibid.*, **95**, 413.
HUBER, P., BALDINGER, E., and HAEBERLI, W., 1950, *Helv. Phys. Acta*, **23**, Suppl. III, 85.
JENTSCHKE, W., 1935, *S.B. Akad. Wiss. Wien*, **144**, IIa, 151 ; 1940, *Phys. Z.*, **41**, 524.
JESSE, W. P., FORSTAT, H., and SADAUSKIS, J., 1950, *Phys. Rev.*, **77**, 782.
JESSE, W. P., and SADAUSKIS, J., 1949 a, *Phys. Rev.*, **75**, 1110 ; 1949 b, *Ibid.*, **76**, 163 ; 1952, *Ibid.*, **88**, 417 ; 1953, *Ibid.*, **90**, 1120 ; 1955, *Ibid.*, **97**, 1668.
JOHNS, H. E., BERNER, J. P., and CORMACK, D. V., 1955, *Nature, Lond.*, **176**, 560.
JOHNSON, J. B., 1917, *Phys. Rev.*, **10**, 609.
KNIPP, J. K., and LING, R. C., 1950, *Phys. Rev.*, **80**, 106 ; 1951, *Ibid.*, **82**, 30.
KULENKAMPFF, H., 1926, *Ann. Phys., Lpz.*, **80**, 91.
LAUGHLIN, J. S., BEATTIE, J. W., HENDERSON, W. J., and HARVEY, R. A., 1953, *Amer. J. Roentgenol.*, **70**, 294.
LEACHMAN, R. B., 1951, *Phys. Rev.*, **83**, 17 ; 1952, *Ibid.*, **87**, 444.
LEHMANN, J. F., 1927, *Proc. Roy. Soc.* A, **115**, 624.
LEHMANN, J. F., and OSGOOD, T. H., 1927, *Proc. Roy. Soc.* A, **115**, 609.
MADSEN, B. S., 1945, *K. Danske Vidensk. Selsk., Mat.-fys. Medd.*, **23**, No. 8.
MEITNER, L., 1926, *Naturwissenschaften*, **14**, 119.
MELTON, C. E., HURST, G. S., and BORTNER, T. E., 1954, *Phys. Rev.*, **96**, 643.
MEYER, S., and SCHWEIDLER, E., 1927, *Radioaktivität*, 2nd ed. (Leipzig : Teubner).
MOTT, N. F., and MASSEY, H. S. W., 1949, *The Theory of Atomic Collisions* (Oxford : Clarendon Press).
NAIDU, R., 1934, *Ann. Chim. (Phys.)*, **1**, 72.
PERRY, W. E., 1953, *Standardisation of Radioactive Isotopes* (National Physical Laboratory).
PIGGE, H., 1934, *Ann. Phys., Lpz.*, **20**, 233.
PONTECORVO, B., 1950, *Helv. Phys. Acta*, **23**, Suppl. III, 97.
RHODES, J., FRANZEN, W., and STEPHENS, W. E., 1952, *Phys. Rev.*, **87**, 141.
SCHMIEDER, K., 1939, *Ann. Phys., Lpz.*, **35**, 445.
SCHMITZ, W., 1928, *Phys. Z.*, **29**, 846.
SELIGER, H. H., and SCHWEBEL, A., 1954, *Nucleonics*, **12**, 54.

SHARPE, J., 1952, *Proc. Roy. Soc.* A, **65**, 860.
STEBLER, A., HUBER, P., and BICHSEL, H., 1949, *Helv. Phys. Acta*, **22**, 362.
STEENBECK, M., 1928, *Ann. Phys., Lpz.*, **87**, 811.
STETTER, G., 1943, *Z. Phys.*, **120**, 839.
TAYLOR, T. S., 1911, *Phil. Mag.*, **21**, 571 ; 1913, *Ibid.*, **26**, 402.
TOWNSEND, J., 1947, *Electrons in Gases* (London : Hutchinson).
TUNNICLIFFE, P. R., and WARD, A. G., 1952, *Proc. Phys. Soc.* A, **65**, 233.
VALENTINE, J. M., 1952, *Proc. Roy. Soc.* A, **211**, 75 ; 1953, *Thesis*, Glasgow University.
VALENTINE, J. M., and CURRAN, S. C., 1952, *Phil. Mag.*, **43**, 964.
WEISS, J., and BERNSTEIN, W., 1955, *Phys. Rev.*, **98**, 1828 ; 1956, *Ibid.*, **103**, 1253.
WERTENSTEIN, 1913, *Thèse*, University of Paris.
WILHELMY, E., 1933, *Z. Phys.*, **83**, 341.
WILLIAMS, E. J., 1932, *Proc. Roy. Soc.* A, **135**, 108.
WILSON, C. T. R., 1923, *Proc. Roy. Soc.* A, **104**, 200.

ABSORPTION AND FLUORESCENCE SPECTRA OF IONS IN CRYSTALS

By W. A. RUNCIMAN
Atomic Energy Research Establishment, Harwell, Berks.

CONTENTS

	PAGE
§ 1. Introduction	30
1.1. Free ion spectra	32
1.2. The static lattice	33
1.3. The adiabatic approximation	34
§ 2. Experimental methods	36
§ 3. Ions with an incomplete d-shell	37
3.1. The octahedral complex	37
3.2. The 3d group	39
3.3. The 4d group	42
3.4. The 5d group	43
3.5. Covalent complexes	43
§ 4. Rare-earth ions	43
4.1. Outline of theory	43
4.2. Polarization spectra	45
4.3. The Zeeman effect	47
4.4. Fluorescence spectra	47
4.5. Trivalent ions	48
4.6. Divalent ions	51
§ 5. Heavy metal ions	52
§ 6. Uranium and the transuranics	53
§ 7. Conclusion	56
Acknowledgments	56
References	56

Abstract. The absorption and fluorescence spectra provide direct information about the energy levels of the ions in a crystal. Alternatively, the ions may be regarded as probes measuring the crystalline field. The introduction deals with the influence of the crystal on the energy levels of the free ions. After a brief description of the experimental techniques, which include measurements down to 1°K, the various ion groups are considered in some detail. Ions of the 3d group are often in approximately octahedral surroundings, and the spectra of such complex ions are described. Then the 4d and 5d groups are discussed. The theory of the spectra of rare-earth salts is reviewed with special reference to polarization and Zeeman effects, and the individual ions are considered. Lastly there are two sections devoted to the heavy metal ions, including thallium, and to uranium and the transuranics, which are of interest because of the importance of the 5f electrons in the interpretation of the spectra.

§ 1. Introduction

The study of the optical spectra of crystals is one of the oldest branches of spectroscopy, and some of the substances found to be of interest in the late nineteenth century, e.g. ruby and the uranyl salts, are still actively investigated today. Since optical absorption is essentially a process in which a *quantum* of light is absorbed, it is not surprising that theoretical understanding was delayed

until quantum mechanics was established in the nineteen-twenties. However, physicists were being attracted to new fields (neutrons have a lot to answer for!), and most of the investigations into luminescent materials of the next twenty years were largely empirical, often stimulated by the growing needs of industry for fluorescent materials. The vast body of experimental material available has been most ably collected and presented by Pringsheim (1949). The luminescence of solids has attracted a lot of attention, and books by Kröger (1948) and Garlick (1949) are also recommended, the latter being particularly useful as a guide to the wider topic of the physical properties, e.g. photoconductivity, of luminescent materials.

This report will attempt to cover the years since the above-mentioned volumes appeared. This is a most significant period, as the results of paramagnetic resonance experiments provided a great stimulus to the understanding of the effects of crystalline and magnetic fields on the energy levels of ions. Many of the same substances are of interest for optical and paramagnetic resonance research, as ions with incomplete electronic shells have a large number of energy levels in the optical region. Paramagnetic resonance techniques can only find low-lying energy levels ($\sim 0{\cdot}1$ cm^{-1}); whereas for optical absorption and fluorescence both low and high energy levels are involved (0 to 50 000 cm^{-1}). Nevertheless the two techniques yield inter-related results as the splitting of the lowest energy levels is often related to the nature and spacing of the excited states. For an account of paramagnetic resonance techniques and results, reference may be made to two earlier reports in this series (Bleaney and Stevens 1953, Bowers and Owen 1955). The corresponding optical data have been tabulated by Hellwege (1955). However, certain substances containing heavy metal ions, e.g. thallium-activated potassium chloride KCl(Tl), and also some compounds containing complex ions, e.g. tungstates, have non-degenerate ground states, and therefore no paramagnetic resonance spectrum, but they can still be studied optically.

The interest in the absorption and fluorescence spectra in solids is two-fold. First it can give information about the energy levels in the ions, and secondly the ions may be considered as probes which can be used to study the crystalline field. Attention will be focused in this review on the former method of approach, as the use of ions as probes depends not only on a knowledge of the energy levels of the ions, but also on their wave functions which have in many cases not yet been calculated. The problem of the interpretation of fluorescence and absorption spectra is a dynamic one as the crystal is vibrating, even if low temperatures are employed to minimize this. Furthermore there may be interaction between absorption centres which will affect the spectra. These effects complicate the analysis of results, but increase the possible scope of optical methods in solid-state research. In order to place the individual problems in perspective, the properties of free ions will be briefly outlined, followed by a discussion of the effects due to the crystal field.

Unfortunately it will not be possible to include an account of irradiation effects. The irradiation colours and luminescence in minerals have been recently described by Przibram (1956). Many interesting problems are raised by the fluorescence of minerals, an example being the origin of the vibrational structure in the sulphur-containing aluminosilicates of the sodalite type (Kirk 1954).

1.1. *Free Ion Spectra*

The energy levels of a free ion, i.e. an ion not embedded in a crystal, are dependent on the Coulomb interaction between electrons, and also on the electronic spin–orbit interaction. If the spin–orbit interaction is small, the Russell–Saunders (LS) scheme is applicable in which energy levels are labelled by the total orbital angular momentum L and the total spin S, in addition to the total angular momentum J, where $\mathbf{J} \equiv \mathbf{L} + \mathbf{S}$. J is always a *good* quantum number for the free ion, since the interactions within the ion only mix states with the same J value.

The total orbital angular momentum is denoted by a letter according to the convention :

$L =$	0	1	2	3	4	5	6	7	8	9	10
	S	P	D	F	G	H	I	K	L	M	N

The spin multiplicity, which is equal to $2S+1$, is written as a prefix and the total angular momentum J as a suffix, e.g. $^3P_2$ stands for a state with $L = 1, S = 1$, $J = 2$. We will not need to be concerned with cases where the spin–orbit interaction dominates and the j–j coupling scheme applies ; but for all the heavier elements there is appreciable spin–orbit interaction. Therefore, although LS labels are useful, detailed calculations have to include spin–orbit interaction in the intermediate coupling approximation. These energy levels schemes are fully described by Condon and Shortley (1935), who give methods for calculating the energy levels.

The ground state of an ion can be rapidly determined by Hund's rules if the electronic configuration is known. These rules state that of all the terms allowed by the Pauli principle, the lowest term will be one of maximum multiplicity. Of those terms with maximum multiplicity, that one with the greatest L value will be lowest. Furthermore for configurations consisting of electrons in a less than half-filled subgroup the spin–orbit splitting is usually normal, i.e. smallest J lowest ; and for configurations with more than a half-filled subgroup the multiplets are usually inverted. As an example, the Pauli principle allows ^{1}SDGI and ^{3}PFH for the configuration f^2, neglecting as always any completely filled subgroups. The ground state in accordance with Hund's rule is $^3H_4$. (Warning : Hund's rules are *not* to be applied to excited states.)

Electric dipole transitions either in absorption or emission can occur for the free ions when the selection rules permit. These are $\Delta J = 0, \pm 1$ (*not* $0 \longleftrightarrow 0$) ; and in addition if LS coupling holds, $\Delta S = 0$, $\Delta L = 0, \pm 1$ (*not* $0 \longleftrightarrow 0$). Laporte's rule states that the parity of the ionic state must also be changed for an allowed transition. The parity is even if the arithmetic sum of the angular momenta of the individual electrons, Σl, is even, and odd if this sum is odd. The half-life of an excited state is about 10^{-8} sec for a free ion if it is capable of an allowed transition to a lower state, the energy difference between states being assumed such that the emitted radiation will be in the visible or near visible spectral range. In inorganic fluorescent materials the half-lives are usually longer than this ($\sim 10^{-7}$–10^{-2} sec), and it is to be expected that the transitions will often be forbidden on the basis of the above selection rules for free ions.

Transitions which are forbidden by the electric dipole selection rules are sometimes observed in free ion spectra. Weak electric quadrupole and magnetic

dipole transitions can occur between states with the same parity. For electric quadrupole transitions, the selection rules are

$$\Delta J \leqslant 2, \quad (not\ 0 \rightarrow 0, 1)$$
$$\Delta L \leqslant 2, \quad (not\ 0 \rightarrow 0, 1).$$

The corresponding selection rules for magnetic dipole transitions are

$$\Delta J \leqslant 1 \quad (not\ 0 \rightarrow 0)$$
$$\Delta L \leqslant 1 \quad (not\ 0 \rightarrow 0).$$

The selection rules on L only apply for Russell–Saunders coupling.

1.2. *The Static Lattice*

Luminescent solids may consist either of pure compounds or of impurity-activated materials. As there are many new features when an ion is embedded in a crystal, first consider the simplified model of a static lattice, in which all vibrational effects are neglected.

The energy levels of the ion will have their $(2J+1)$-fold degeneracy either partially or completely removed by Stark splitting in the electric crystalline field. Now all the energy levels will have new crystal quantum numbers which can be used to label the states. These crystal quantum numbers will be associated with the irreducible representations belonging to the point group of the ion site in the crystal. For instance, if the crystal field is strong, the orbital angular momentum L will be *quenched*, and hence will cease to be a good quantum number for labelling states. Nevertheless, the total spin S may still be a good quantum number, as the spin–lattice interaction (Bleaney and Stevens 1953) is usually small.

The transitions observed in the solid state are often between states of the same parity. These could be, but very rarely are, transitions of the magnetic dipole or electric quadrupole type. Instead, in a crystal field which has no centre of symmetry the state of an electronic configuration will get mixed with states of different parity, and this will enable electric dipole transitions to take place even between two states of the same nominal configuration. Such transitions are sometimes called forced electric dipole. For ions at a centre of symmetry, asymmetric vibrations can enable electric dipole transitions to occur (Van Vleck 1937).

The site symmetry of the ion may be reduced from that in the pure crystal if the ion is present as an impurity. Firstly, if the ion has a degenerate ground state in the crystal field, then the Jahn–Teller theorem asserts that the lattice configuration is unstable and the lattice will distort, thus lowering the symmetry in the neighbourhood of the ion. Secondly, if an impurity ion is of different valency from the ion which it replaces in the pure crystal, then the symmetry may be lowered by associated defects locally compensating the excess charge.

When an ion in a pure crystal is highly excited it may lose an electron to another ion. This is the cause of the charge transfer spectra which lead to strong absorption in the ultra-violet. If all the ions in the crystal have an inert gas configuration in the ground state, then the solid will have no visible absorption and this fundamental absorption will occur at wavelengths shorter than the absorption edge, ~ 2000Å. Such materials are obviously good base materials for incorporating impurity ions which do have a detailed absorption in the visible range, 4000–7000Å.

Owing to the high density of solids as compared with gases, even forbidden transitions can cause appreciable absorption when impurity ions are incorporated in solids.

When an electron is excited to such an extent that it is freed from its ion, it is normally in a conducting state; but in some compounds it is possible for the electron to travel through the lattice bound to a compensating *hole*. This combination is called an *exciton*. The absorption and fluorescence spectra of excitons have much in common with those of ions, but always occur just to the long wavelength side of the absorption edge. Cu_2O and HgI_2 have well-defined exciton spectra, but it is still not clearly demonstrated that the spectra are independent of impurities in the crystal.

1.3. *The Adiabatic Approximation*

The above model of an ion embedded in a static lattice is useful if the electronic transition to be considered does not change the surroundings of the ion. This is especially true for the rare-earth ions, where the 4f electrons responsible for the absorption and emission spectra are shielded from the neighbouring ions by 5s and 5p electrons. In these circumstances the absorption energy which raises the ion to its lowest excited states is the same as the fluorescence energy from these states. This leads to well-defined lines in the absorption and fluorescence spectra.

On the other hand, when the outer electrons are involved in an electronic transition, there will usually be a change in the effective radius of the ion. The transition will induce lattice vibrations which will result either in a vibrational structure being observable in the spectra or else in the appearance of very broad bands. The dissipation of some of the absorbed energy thermally accounts for Stokes' law which states that fluorescent emission occurs at wavelengths equal to or longer than the absorption wavelength.

A full treatment of the optical excitation of an impurity ion in a crystal is so complex that it is fortunate that approximations can often be usefully employed. The adiabatic or Born–Oppenheimer approximation depends on the small ratio of the electronic to nuclear mass, and uses this to separate out the calculation of electronic and nuclear motions. In this approximation the nuclei can be considered as vibrating in the averaged potential of the electrons. The adiabatic approximation is valid so long as the spacing between the electronic energy levels is large compared with the vibrational energies, since otherwise the vibrations cause a mixing of the electronic states. Hence if an electronic energy level is degenerate, the adiabatic approximation will break down. This will be the case if an ion which is a constituent in a pure solid is excited, since the energy is the same as that obtained if any of the other ions were excited. This may lead to an increase in the number of lines in the absorption spectrum. The splittings (Davydov splitting) so produced, though large in some organic crystals, are usually negligible in inorganic crystals due to the forbidden nature of the electronic transitions involved. Unresolved splittings can however contribute to the line widths of absorption and fluorescent lines.

The adiabatic approximation is the basis for the configuration coordinate model. The positions of the nuclei are represented by a single variable x, e.g. the distance of the nearest neighbours from the impurity ion. The potential energy of the nuclei in the ground electronic state is shown as a function of x by

curve FAG in figure 1. Similarly in the excited state the nuclei move in a potential well FCH. In the harmonic approximation the potential energy curves are parabolic. The energy levels of the centre are represented by horizontal lines the lowest of which is a little above the minimum of the parabola on account of the zero-point energy. The higher lines refer to states in which some vibrational quanta are present. The wave function of the nuclear motion is approximately Gaussian for the lowest state, and has the largest values near the parabola for a highly excited state as shown in figure 1.

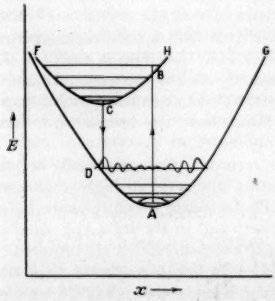


Figure 1. Configuration coordinate diagram.

The Franck–Condon principle states that during an electronic transition the positions and momenta of the ions will be substantially unchanged, owing to the large ratio of ionic masses to the electronic mass. Hence on figure 1 electronic transitions are shown as vertical lines, AB for absorption and CD for emission. After the absorption of light AB, the centre will relax BC to the new equilibrium position C, some of the vibrational energy being dissipated thermally to the rest of the lattice. Similarly heat is dissipated after fluorescence CD, while the centre relaxes to its original state. It follows that the fluorescence is normally at longer wavelengths than the absorption.

This configuration coordinate diagram has been used by Dexter, Klick and Russell (1955) to formulate a criterion for luminescence. If the intersection of the parabolae at F is lower than the point B reached on the upper parabola when energy is absorbed, then they suggest that the vibrations are sufficiently undamped for the configuration to cross to the other side of the parabola near F. It will then transfer to the lower parabola without emitting radiation before equilibrium at C can be reached.

This suggestion seems valid provided restrictions due to spin or spatial symmetry do not prevent the radiationless cross-over at F. The situation is obscure near F due to the breakdown of the adiabatic approximation whenever electronic levels are near-degenerate.

In general, fluorescence is more probable for a large energy gap between the upper and lower states, as the transition probability for radiation is proportional

to the cube of the radiation frequency, whereas radiationless transitions decrease in probability, as more vibrational quanta are needed to dissipate the energy. A spin change may only be effected by lattice vibrations through the spin–orbit interaction, hence there may be a rule for inorganic materials that the fluorescent state, if one exists for the system in question, is the lowest state with spin differing from the ground state. Fluorescence may be absent if this excited state is heavily mixed with a state whose spin is the same as that of the ground state.

§2. Experimental Methods

The absorption spectrum of a single crystal is obtained by analysing a light beam, originating from a source with a continuous energy spectrum, after it has traversed the crystal. Black body radiation is useful in the visible and infra-red, and the source most commonly used is the 6 volt, 18 amp, ribbon-filament lamp. For wavelengths below 3400 Å a hydrogen lamp with a quartz window is necessary. Xenon arcs and zirconia lamps have also been used to obtain continuous spectra, but are more difficult to operate.

The spectra are either recorded photographically or photoelectrically. As the sources are of only moderate intensity, all high-resolution work has so far been done photographically. Photoelectric measurements have been obtained in the visible using photomultipliers and in the infra-red using lead-sulphide cells, and are to be preferred if intensity measurements are required. Glass or quartz prism spectrographs have been used for low or moderate dispersion,† and grating instruments for high dispersion, e.g. Dieke and Singh (1956) obtain a first-order dispersion of 1·2 Å mm^{-1} using a 21 ft concave grating spectrograph. Wavelength measurements are made by comparison with a standard source, e.g. an iron arc.

If crystals have less than cubic symmetry, further information can be gained by inserting a polarizer into the light path. A Nicol calcite polarizer can be used, or a quartz double prism of the Wollaston type if both polarization spectra are to be photographed simultaneously. For wavelengths below 3000Å, the components are usually joined with glycerin or castor oil rather than Canada balsam; but gedamine provides an adhesive useful to 2500Å (Bouriau and Lenoble 1957). Cubic crystals can be studied using polarized light if the intensity of the radiation is measured in different directions (Feofilov 1956 a).

For Zeeman work, a magnetic field of 25–40 kilogauss is necessary, as, owing to the natural line width, it is not possible to resolve the components with smaller fields. A pole gap of about 1 inch is needed to accommodate the crystal and cryostat. A typical iron-core magnet satisfying these requirements weighs over two tons, and dissipates 18 kw of power (Bowen 1957).

Most of the apparatus is the same for fluorescence, except that instead of irradiating with a continuous source, one irradiates preferentially in a wavelength region where there is a high excitation probability. Low pressure and high pressure mercury vapour lamps produce short wavelength (2537Å) and long wavelength (3650Å) ultra-violet radiation respectively. The high pressure lamp can be obtained with a coloured glass outer bulb, then known as a *black* lamp, to filter

† Quartz optics can give absorption bands a spurious banded structure, due to the optical rotation of the quartz (Ellis and Kaplan 1935, 1937). This appears if the incident light is polarized either by the dichroism of a crystal or by the use of a conventional polarizer.

out the visible radiation. If required, the residual near infra-red radiation it emits can be filtered out with a copper sulphate solution (125g $CuSO_4.5H_2O$ per litre of water) as recommended by Bowen (1946).

The crystal under investigation is often cooled to liquid oxygen (90°K), nitrogen (77°K), hydrogen (20°K) or helium (4·2°K) temperatures. In the case of liquid helium, a double Dewar is used, the outer containing liquid nitrogen or oxygen. Temperatures down to 1°K may be obtained by keeping the helium at reduced pressure.

If single crystals are not available the spectra from powder samples can yield valuable information, the fluorescence spectrum being obtained normally, and the absorption spectrum being replaced by a similar reflection spectrum. However, it is not usually possible to obtain polarization or Zeeman spectra from powders. For work at low temperatures, the powders can be inserted in glass or quartz tubes, containing helium at a few millimetres of mercury pressure to remove heat from the powder. Single crystals are grown from the melt or from solution, whereas activators are often incorporated in powders by heating in a furnace, sometimes in a controlled atmosphere.

§3. Ions with an Incomplete d-shell

3.1. *The Octahedral Complex*

Ions of the 3d, 4d and 5d groups often occupy sites of six-fold coordination, the surrounding ions being arranged in nearly octahedral symmetry. The crystal field is strong enough to perturb and sometimes to dominate the effect of the Coulomb forces in the ion. In a strong field the five magnetic substates of a d-electron, each double because of spin, are split into a triplet capable of accommodating six t_{2g} electrons and a doublet for four e_g electrons. The separation between the possible energy levels will be denoted by Δ (Owen 1955) and replaces the older $10Dq$. In octahedral surroundings the triplet is lower than the doublet, whereas the reverse is true for a tetrahedral arrangement of negative ions around the central ion. These t_{2g} and e_g orbitals are described by Bowers and Owen (1955), who also give a brief description of the crystallography of many transition metal salts. Hush and Pryce (1957) have shown that irregularities in the ionic radii of iron group ions are due to the different occupation of the two types of orbital.

Tanabe and Sugano (1954 a, b) and Orgel (1952, 1955 a, b, c, d) have calculated the effect of a cubic field on the energy levels of ions with an electronic configuration d^n. The results are best shown graphically, and as an example the energy level diagram for d^3 is shown in figure 2. The energy ordinate E and the strength of the crystal field Δ are both in units of B, a linear combination of Slater integrals introduced by Racah. At the left-hand side are the Russell–Saunders levels and at the right are the values for a strong field. Irrespective of the crystal field strength, the spin and the symmetry are good quantum numbers. The symmetries are labelled by the irreducible representations of the cubic group A_1, A_2, E, T_1 and T_2, where A is singly, E doubly and T triply degenerate.

A strong crystal field will sometimes alter the spin and spatial symmetry of the ground state, but not for d^3, figure 2, which always has a $^4A_2$ ground state

independent of field strength. Since energies are measured relative to the ground state this change in the character of the ground state, when it occurs, causes a discontinuity in the slope of the energy levels in the diagram. In weak or medium fields the spin of the ground state is the maximum possible in accordance with Hund's rules ; but in strong fields the spin is only the maximum consistent with first filling all the t_{2g} states before any e_g states.

The octahedral surroundings of the ion will be distorted by the Jahn–Teller effect if the ground state of the ion in a cubic field is degenerate. This effect has been studied recently by Öpik and Pryce (1957) and by Orgel and Dunitz (1957), and has been shown to be large in some cases, especially for cupric salts.

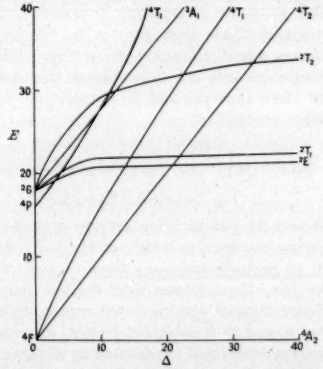


Figure 2. Energy diagram for d^3 in a cubic field.

However, there may also be considerable distortions of the octahedral surroundings due to more general crystallographic considerations. These distortions and also spin–orbit interaction will split the levels which are degenerate in the cubic field. The spin–orbit splitting will have a first-order effect on levels of symmetry types T_{1g} and T_{2g}. Second-order effects will usually remove all degeneracy, except that for ions with an odd number of electrons the levels will consist of Kramer's doublets even in fields of lowest symmetry. Some aspects of ions containing d-electrons in fields of less than cubic symmetry have been theoretically investigated by Ballhausen and Jørgensen (1955). For a more general account of crystal field theory applied to d-electrons reference may be made to the review of Moffitt and Ballhausen (1956).

The line-width of an absorption line will be small if the energy level of the excited state is parallel or near parallel to the energy level of the ground state in the energy level diagram. This condition obtains for sufficiently strong crystal fields when the number of t_{2g} and e_g electrons is the same for both the ground and excited state. For example, in figure 2 this is true for the states ^2E, $^2T_1$ and $^2T_2$ which have the same configuration t_{2g}^3 as the ground state $^4A_2$. The ionic radius

is not then appreciably affected by the transition, and hence the transition will not excite many vibrations in the lattice. Furthermore the thermal and zero point energy of the lattice will cause a variation in the effective crystal field strength when the electronic transition takes place, and, if the lines in the energy level diagram are not parallel, this will lead to a broadening of the absorption line. It follows from this criterion that transitions from the ground state which involve no spin change are always broad. They are also the strongest transitions, since they are the least forbidden, and may obscure any narrow lines due to intersystem transitions. Ballhausen (1955) and Liehr and Ballhausen (1957) have considered the effect of symmetry on the intensities of the absorption bands of octahedral complexes.

When detailed calculations are made of the splitting of energy levels in the crystal, the parameters have to be modified from those of the free ion to include the effects of covalent bonding. Owen (1955) has given a detailed discussion for some hydrated salts. This generalized type of crystal field theory is now usually termed ' ligand field theory '.

3.2. *The* 3d *Group*

The various ions will now be discussed in order of the number of d-electrons which they contain. The hydrated ions have been especially studied, and most recently by Holmes and McClure (1957). After each ion will be listed the ground state, of the ion in an octahedral complex, appropriate to the weak field approximation, which applies moderately well for the hydrated salts. The splitting of the lowest *LS* state is inverted in the d^{10-n} configuration from that of d^n, and hence the ground states have the same spin, but different spatial symmetry.

d^1 : Ti^{3+}, *ground state* $^2T_2$

In a cubic field the only *LS* state ^2D splits into the ground state $^2T_2$ and an excited state ^2E. Hartmann and Schläfer (1951 a) report a single broad band peaked at 20 300 cm^{-1} which can be attributed to the transition $^2T_2 \rightarrow {}^2E$.

d^2 : V^{3+}, *ground state* $^3T_1$

Hartmann and Schläfer (1951 b) report bands at 17 800 and 25 600 cm^{-1}, which can be assigned to the transitions $^3T_1 \rightarrow {}^3T_2$ and $^3T_1 \rightarrow {}^3T_1$, as these are the transitions most likely to give rise to strong absorption bands.†

d^3 : V^{2+}, Cr^{3+}, Mn^{4+}, *ground state* $^4A_2$

The absorption spectrum can be interpreted in terms of the energy level diagram, figure 2. The main absorption bands of hydrated chromium salts are at 17 500 and 24 500 cm^{-1} and are due to transitions to $^4T_2$ and $^4T_1$ states respectively, transitions to the doublet states being weak as discussed above. Holmes and McClure (1957) have found similar bands for vanadous solutions at 11 800 and 17 500 cm^{-1}. A further band in the chromium salts at 36 600 cm^{-1} is probably due to a transition to a higher $^4T_1$ state. There are also sharp lines in the absorption

† Low (1957 a) has described a vibrational structure in the absorption band due to the transition $^2T_1 \rightarrow {}^2T_2$ and found an absorption line at 21 000 cm^{-1} in Al_2O_3(v). The author (unpublished) has resolved the line $^2T_1 \rightarrow {}^1A_1$ into two components, differently polarized, and found two further lines in the near infra-red, $^2T_1 \rightarrow {}^1E$, $^1T_2$.

spectra of chromium salts caused by transitions to the doublet states ^2E, $^2T_1$ and $^2T_2$. Spedding and Nutting (1934, 1935) found that in a magnetic field each of the three strong absorption lines of potassium chrome alum, $KCr(SO_4)_2.12H_2O$, splits into three elliptically polarized components with a separation close to the 'spin only' value, i.e. $g = 2$. The theoretical interpretation is not complete, but

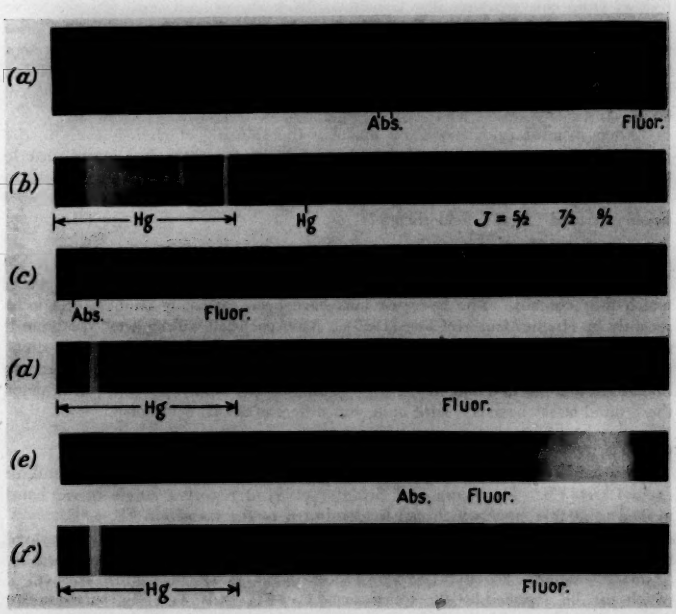


Figure 3. Crystal spectra obtained using a medium quartz spectrograph. The scale is in units of 100Å. (*a*) Absorption spectrum of ruby at room temperature. (*b*) Fluorescence spectrum of CaO(U, Sm) at 77°K. (*c*) Absorption spectrum of CaO(Bi) at 77°K. (*d*) Fluorescence spectrum of uranyl nitrate hexahydrate at 77°K. (*e*) Absorption spectrum of uranyl nitrate hexahydrate at 77°K. (*f*) Fluorescence spectrum of NaF(U, O) at 77°K.

calculations by Van Vleck (1940) and by Finkelstein and Van Vleck (1940) demonstrate that the upper state is a doublet. Couture, Jacquinot and Tsujikawa (1956) have found complicated changes in the absorption spectrum of ammonium chrome alum at the transition temperature, *c.* 80°K, and Couture (1957) finds similar effects in potassium chrome alum at 60°K. Low (1957 b) has described the paramagnetic resonance and optical spectra of Cr^{3+} in MgO.

There have been many investigations of the absorption and fluorescence spectra of ruby, $Al_2O_3(Cr)$. Two lines at 14 450 and 14 420 cm^{-1} appear in both

fluorescence and absorption. The upper energy levels in these transitions are two components of the ^2E state, split by the combined effect of spin–orbit interaction and a trigonal component of the crystal field (Tanabe and Sugano 1957). Krishnan (1947) has compared the frequency shifts of vibrational satellites to the above lines with infra-red and Raman frequencies. Deutschbein (1939) has studied the fluorescence of chromium-activated solids down to 1·6°K and has shown that the levels are populated in accordance with a Boltzmann distribution. Slight shifts ($\sim 0{\cdot}1$ cm^{-1}) have been detected by Deutschbein, Joos and Teltow (1942) between the centres of the absorption and emission lines.

The absorption lines between 4600 and 4800Å (figure 3 (*a*)) in ruby are attributable to transitions to $^2T_2$ levels. The polarization of these and the other absorption and fluorescence lines has been studied by Mani (1942). Grechushnikov and Feofilov (1955) have observed a vibrational fine structure in the main absorption band between 5500 and 6000Å when ruby is cooled to liquid nitrogen temperatures.

Prener (1953) has reported a low temperature fluorescence of MgO (Mn, Li). The fluorescence of the tetravalent manganese is analogous to that in chromium-activated solids.

d^4 : Mn^{3+}, Cr^{2+}, *ground state* ^5E

Hartmann and Schläfer (1951 b) find a single absorption band at 21 000 cm^{-1} in $CsMn(SO_4)_2.12H_2O$. Similarly there is an absorption band at 13 900 in $CrSO_4.7H_2O$. Both these may be assigned to the transition $^5E \rightarrow {}^5T_2$.

d^5 : Mn^{2+}, Fe^{3+}, *ground state* $^6A_1$

The ground state is the only sextet state, d^5 being a half-filled shell, and hence it is to be expected that all absorption bands will be weak. Because of this weakness, Holmes and McClure (1957) investigated the spectra of the ions in solution, finding bands at 18 900 and 23 000 cm^{-1} for Mn^{2+} solutions and bands at 12 500 and 18 200 cm^{-1} for Fe^{3+} solutions. These correspond to transitions to the $^4T_1$ and $^4T_2$ states respectively for each ion. Gielessen (1935) observed line spectra in solid $MnCl_2.4H_2O$, and Pappalardo (1957 b) has attributed line groups in manganese salts to transitions to the upper states as follows: 25 000 ($^4A_1$ and ^4E_a), 27 000 ($^4T_2$) and 29 000 cm^{-1} (^4E_b).

Divalent manganese is an extremely efficient luminescent activator in a wide variety of chemical systems. In cubic surroundings the upper state for fluorescent transitions belongs to the $^4T_1$ representation. The fluorescence spectrum consists of a single band, often in the green or orange, which is appreciably asymmetric. As has been pointed out by Klick and Schulman (1952), this asymmetry can arise if the parabolae of figure 1 are only slightly displaced from each other in the horizontal direction.

The configuration d^5 consisting of a half-filled shell has a ground state which is only split because of high-order perturbations. The resultant small splittings have been investigated by paramagnetic resonance (Bleaney and Stevens 1953, Bowers and Owen 1955).

d^6 : Fe^{2+}, *ground state* $^5T_2$

Holmes and McClure (1957) report a band at 10 000 cm^{-1} in $FeSO_4.7H_2O$ which is just resolved into a double peak. Dreisch and Kallscheuer (1944) had

earlier obtained similar results from ferrous chloride solutions, and in each case the bands are probably due to $^5T_2 \rightarrow {}^5E$ transitions.

$d^7 : Co^{2+}$, *ground state* $^4T_1$

Holmes and McClure (1957) find an absorption band at 8350 cm^{-1} due to the transition $^4T_1 \rightarrow {}^4T_2$, and also another band at 19 800 cm^{-1}, resolvable into four components at liquid nitrogen temperature, which they attribute to a superposition of the transitions $^4T_1 \rightarrow {}^4A_2$ and $^4T_1 \rightarrow {}^4T_1$. Gielessen (1935) has studied a wide variety of cobaltous salts at 84°K and found a detailed line spectrum which varies appreciably from salt to salt, and is possibly due to transitions to the doublet states $^2T_1$ and $^2T_2$. Low (to be published) has reported optical and paramagnetic resonance spectra of divalent cobalt in single crystals of MgO and has found no appreciable Jahn–Teller distortions from octahedral symmetry.

$d^8 : Ni^{2+}$, *ground state* $^3A_2$

$NiSO_4.7H_2O$ has three absorption bands at 8600, 14 700 and 25 500 cm^{-1} which are attributed to transitions to the upper states $^3T_2$, $^3T_{1a}$ and $^3T_{1b}$ respectively by Holmes and McClure (1957), in accordance with an earlier analysis of the experimental results by Tanabe and Sugano (1954 b). Gielessen (1935) found a detailed line spectrum in $NiCl_2.6H_2O$ at 84°K in the frequency range 17 000–21 000 cm^{-1}. Pappalardo (1957 a) has observed that the spectra of nickel salts change on cooling from 77°K to either 4°K or 20°K. He attributes line groups near 16 000 and 24 000 cm^{-1} to transitions to ^1E and $^1A_1$ levels respectively. Low (to be published) has obtained and analysed the spectrum of divalent nickel in single cystals of MgO.

$d^9 : Cu^{2+}$, *ground state* ^2E

The single absorption band at 12 200 cm^{-1} in $CuSO_4.5H_2O$ is due to the only transition within the d^9 configuration, $^2E \rightarrow {}^2T_2$. The two-fold degenerate ground state leads to a complicated Jahn–Teller distortion which has been discussed by Öpik and Pryce (1957) and by Orgel and Dunitz (1957). The optical spectra of the deformed complexes have been described by Bjerrum, Ballhausen and Jørgensen (1954), and the theory further developed by Ballhausen (1954).

3.3. *The* 4d *Group*

This group has not been so extensively studied as the 3d group ; but two main classes have received considerable attention.

(*a*) *Molybdates.*

There is a full discussion of the absorption and fluorescence of the molybdates in Kröger (1948). For example, $CaMoO_4$ at 180°c has a broad absorption band whose low energy edge is about 34 000 cm^{-1} and has a fluorescent band at 18 700 cm^{-1}.

(*b*) *Palladium salts.*

Yamada (1951) has shown that the absorption of the planar complexes in K_2PdCl_4 and other salts has a marked dichroism, which he uses in a discussion of the origin of the absorption bands.

3.4. *The* 5d *Group*

This group is in many ways similar to the 4d group, and discussion will be restricted to two main classes as before.

(*a*) *Tungstates.*

The tungstates are also described in detail by Kröger (1948). $CaWO_4$ at $-180°$c has an absorption edge at 41 000 cm^{-1} and an emission peak at 23 800 cm^{-1}. Hameka and Vlam (1953) have derived the configuration coordinate diagram suitable for calcium tungstate.

(*b*) *Platinum salts.*

Yamada (1951) has investigated the dichroism of small crystals of K_2PtCl_4 and other salts with planar complexes.

3.5. *Covalent Complexes*

Some salts, especially the cyanides, have ground states with spins in accordance with the strong field approximation. The t_{2g} orbitals fill first, and with six d-electrons this sub-group is complete. Hence for the configuration d^6 the ground state is $^1A_1$ and the salt is diamagnetic, e.g. $K_3Co(CN)_6$. No salts of this strong field type are known for ions with more than six d-electrons. Owen (1955) has found it necessary to postulate a covalent character in the cyanides to explain the experimental results quantitatively.

The manganate $(MnO_4)^{2-}$, permanganate $(MnO_4)^-$ and chromate $(CrO_4)^{2-}$ ions have complex spectra (Teltow 1938, 1939), especially when present in mixed crystals, e.g. $K_2(S, Mn)O_4$ and $K(Cl, Mn)O_4$. The electronic structure and spectra of the tetrahedral permanganate and chromate ions have been theoretically studied by Wolfsberg and Helmholz (1952).

§ 4. Rare-earth Ions

4.1. *Outline of Theory*

The trivalent rare-earth ions have long been known to have absorption spectra, and in some cases fluorescence spectra, consisting of groups of lines. References to the earlier literature will be found in the books by Pringsheim (1949) and by Yost, Russell and Garner (1947). A single level in the free ion, characterized by total angular momentum J, is split by the crystal field (Stark effect) to produce a group of levels. Each group of lines corresponds to transitions to such a group of levels. The number of levels into which a given level of the free ion is split by the crystal field can be derived by the group-theoretical methods of Bethe (1929). The results for the 32 crystallographic groups have been summarized (Runciman 1956 a). The point groups are classified under four headings : (*a*) cubic groups, (*b*) hexagonal and rhombohedral, (*c*) tetragonal, and (*d*) orthorhombic and lower symmetry groups. The levels are then split as shown in table 1.

Judd (1957 a) has found it helpful in some cases to consider the splitting which would occur in icosahedral symmetry and then to find the effects of the deviations from this idealization. From a detailed consideration of the wave functions, Hellwege (1948) has derived selection rules for absorption by polarized light. Unfortunately these selection rules are often apparently violated because of effects

due to the lattice vibrations. The vibration–electronic interaction has been considered by Satten (1957), with special reference to the spectra of neodymium salts. Because of this difficulty in interpreting the crystal spectra, Elliott, Judd and Runciman (1957) have calculated the energy levels of the multiplets in the Russell–Saunders approximation using the tensor-operator and group-theoretical methods of Racah (1949). Assignments of levels to the various line groups can be made, and tested by detailed calculations of the crystal splittings.

Table 1. Number of Levels produced by Stark Splitting of an Ion Level with Angular Momentum J in Crystal Fields of Different Symmetries

J	0	1	2	3	4	5	6	1/2	3/2	5/2	7/2	9/2	11/2
cubic	1	1	2	3	4	4	6	1	1	2	3	3	4
hexagonal	1	2	3	5	6	7	9	1	2	3	4	5	6
tetragonal	1	2	4	5	7	8	10	1	2	3	4	5	6
orthorhombic	1	3	5	7	9	11	13	1	2	3	4	5	6

The Coulomb energy of the free ion is evaluated in terms of the Slater integrals F_2, F_4 and F_6 as defined in Condon and Shortley (1935). A hydrogenic type wave function has sometimes been used (Judd 1955 b) to determine the ratios of the Slater integrals for a 4f electron, i.e.

$$F_4/F_2 = 0{\cdot}138, \quad F_6/F_2 = 0{\cdot}0151.$$

Then the Coulomb energy levels can be expressed in terms of the single parameter F_2, which in turn can be expressed as a function of the atomic number Z of the rare-earth ion by the equation (Judd 1956)

$$F_2 = 12{\cdot}4(Z-34)\ \mathrm{cm}^{-1}.$$

When the Coulomb energies are expressed in terms of F_2, the same energy level diagram holds for f^n and f^{14-n}. The lowest levels determined by Elliott, Judd and Runciman (1957) are shown in figure 4. These values are in some cases different from those found earlier by Jørgensen (1955 a), who incorporated fewer states in his calculations and used different ratios for the Slater integrals. It must be emphasized that figure 4 only gives a general indication of the energy levels, and it will be necessary to introduce appreciable spin–orbit coupling in order to get good agreement with experiment. Also it is to be hoped that calculated wave functions for the rare-earth ions will become available to avoid the necessity of a rather arbitrary choice for the Slater integrals.†

A weak spin-orbit interaction splits each of the Coulomb levels into a multiplet. The shift of a component of a multiplet from the centre of the multiplet is given by

$$\Delta\nu = \tfrac{1}{2}\lambda\xi[J(J+1)-L(L+1)-S(S+1)],$$

where λ is the spin–orbit splitting factor and ξ is the spin–orbit interaction for a single 4f electron. Bleaney (1955) has derived semi-empirical values of ξ for all rare-earth ions. The spin–orbit splitting factors are listed in Elliott, Judd and Runciman (1957) for all levels for which the Coulomb energies have been calculated,

† Dr. E. C. Ridley (unpublished) has just completed her calculations of the wave functions for Pr^{3+} and Tm^{3+} using the method of self-consistent fields, without exchange.

i.e. all levels of f^2, f^3 and f^4 and the two highest multiplicities of f^5, f^6 and f^7. The spin–orbit splitting factors for f^{14-n} are the same as those for f^n, but with reversed sign. This paper also describes how to calculate the matrix elements of the spin–orbit matrices, if it is found necessary to use intermediate coupling calculations which allow mixing of LS states.

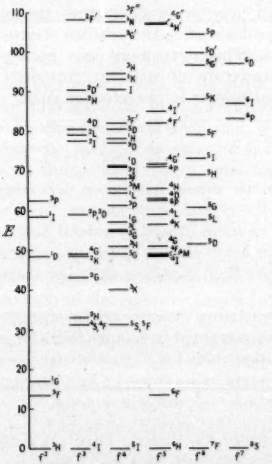


Figure 4. Energy levels of means of multiplets for f^n.

Owing to the screening of the 4f electrons by the outer electron of the rare-earth ion, the effect of the crystal field can be considered to be small so that normally we may keep J as a quantum number and calculate crystal field splittings by the methods of perturbation theory. Most of the salts which have been extensively studied have trigonal symmetry. In the ethylsulphates the rare-earth ion is on a site of C_{3h} symmetry for which the potential energy of an electron may be written

$$V = A_2^0 V_2^0 + A_4^0 V_4^0 + A_6^0 V_6^0 + A_6^6 V_6^6,$$

where the V_k^q transform like spherical harmonics Y_{kq}. The axial terms V_k^0 cause a splitting of levels with differing $|M|$, where M is the z-component of J, whereas V_k^q mixes states for which $M - M' = q$. Hence it is convenient to define a crystal quantum number μ, where $M \equiv \mu \pmod{q}$.

4.2. *Polarization Spectra*

The only transitions which are likely to occur in rare-earth spectra are electric dipole, electric quadrupole or magnetic dipole. These can be partially resolved as follows. The absorption spectrum of a uniaxial crystal is measured with the

light path parallel to the crystallographic axis for the ' axial ' spectrum, and with the light perpendicular to the axis for the transverse spectrum. The transverse spectrum is split into π and σ spectra depending on whether the axis of polarization, i.e. the direction of the electric vector, is parallel to or perpendicular to the crystal axis. If a line in the axial and σ spectra coincide the transition is electric dipole, and if the axial and π spectra coincide the transition is magnetic dipole or electric quadrupole. It is often possible to distinguish the magnetic dipole from the electric quadrupole transitions using the detailed selection rules (Hellwege 1948), and so far only magnetic dipole transitions have been positively identified, these being present in the spectrum of trivalent europium salts. Magnetic dipole transitions can only occur subject to the selection rules,

$$\Delta J \leqslant 1, \quad (not\ 0 \rightarrow 0)$$
$$\Delta L \leqslant 1, \quad (not\ 0 \rightarrow 0).$$

On the other hand, electric dipole transitions occurring because of the lack of centre of symmetry in the crystal allow changes in J and L up to seven units, owing to the admixture of states caused by crystal field terms up to order six. A rather curious feature for some electric quadrupole transitions is that they might be present for an arbitrary light direction although absent in both transverse and axial spectra.

In view of the overwhelming majority of the transitions being electric dipole, the selection rules will be presented in tabular form for two common symmetries. Table 2 shows the selection rules for C_{3h} symmetry between states characterized by the crystal quantum number $\mu_I \equiv \mu \pm \frac{1}{2}p\,\Sigma_k l_k$ (mod p), where the sum is over the f-electrons which have $l_k = 3$ and where $p = 6$ for C_{3h} symmetry.

Table 2. Selection Rules for Electric Dipole Transitions, C_{3h} Symmetry

(*a*) Even number of electrons.

μ_I	0	±1	±2	3
0			σ	π
±1		σ	π	σ
±2	σ	π	σ	
3	π	σ		

(*b*) Odd number of electrons.

μ_I	±1/2	±3/2	±5/2
±1/2		σ	σ,π
±3/2	σ	π	σ
±5/2	σ,π	σ	

For an even number of electrons, the π and σ spectra are completely different. Satten (1957) has shown that this dichroism is preserved when vibrational interactions are included. Examples where C_{3h} symmetry applies are rare-earth impurities present in anhydrous lanthanum chloride, $LaCl_3$, and rare-earth ethylsulphates, e.g. $Nd(C_2H_5SO_4)_3.9H_2O$.

Another common symmetry type is C_{3v} (Judd 1955 a, c), which includes the double nitrates, e.g. $Pr_2Mg_3(NO_3)_{12}.24H_2O$, and the bromates, e.g. $Nd(BrO_3)_3.9H_2O$. The selection rules are as in table 3, where the auxiliary quantum number S distinguishes those states symmetrical or anti-symmetrical with respect to reflection in the vertical reflection planes.

The corresponding table for C_3 symmetry is simply obtained by dropping the auxiliary quantum number S, and hence allowing all transitions between $\mu = 0$ states with π polarization.

Table 3. Selection Rules for Electric Dipole Transitions, C_{3v} Symmetry

(*a*) Even number of electrons.

μ		0		± 1
		$S=1$	$S=-1$	
0	$S=1$	π		σ
	$S=-1$		π	
± 1		σ		σ, π

(*b*) Odd number of electrons.

μ	$\pm 1/2$	$\pm 3/2$
$\pm 1/2$	σ, π	σ
$\pm 3/2$	σ	π

4.3. *The Zeeman Effect*

For an ion which is not in a crystal field, a magnetic field introduces a term into the Hamiltonian of the form $\beta(\mathbf{L}+2\mathbf{S}).\mathbf{H} \equiv \beta(\mathbf{J}+\mathbf{S}).\mathbf{H}$ which has eigenvalues $g\beta HM$ where β is the Bohr magneton, $\equiv eh/4\pi mc$, and M is the component of $\mathbf{J}$ in the direction of the magnetic field. The value of the Landé factor g in the Russell–Saunders scheme is independent of all quantum numbers other than S, L and J and is given by

$$g = 1+\frac{J(J+1)-L(L+1)+S(S+1)}{2J(J+1)}.$$

In the intermediate coupling approximation J is still a good quantum number, and g is simply a weighted sum of the separate Landé factors for the constituent LS states (Condon and Shortley 1935).

The original theory of the Zeeman effect in crystals (Bethe 1930) has been elaborated for all 32 point groups by Hellwege (1950 a, b). If in a uniaxial crystal the magnetic field is applied parallel to the crystal axis each doublet will be split by an amount $s_1\beta H$ where the splitting factor s_1 introduced by Dieke and Heroux (1956) is the weighted sum of the splitting factors $2gM_i$ of the constituent states, i.e.

$$s_1 = 2g(a_1^2 M_1 + a_2^2 M_2 + \ldots).$$

If the axial part of the field dominates, and the magnetic field is perpendicular to the crystal axis, then for an odd number of electrons only the $M = \pm\frac{1}{2}$ level is split and that by an amount $g(J+\frac{1}{2})\beta H$.

In general, the spectrum may be complex as the crystal field may be strong enough to mix states of different J. Such effects have been included in the theory of the magnetic splittings of the ground states which have been studied by paramagnetic resonance (Bleaney and Stevens 1953, Bowers and Owen 1955, Stevens 1952, Elliott and Stevens 1952, 1953 a, b).

4.4. *Fluorescence Spectra*

Rare-earth ions are frequently fluorescent when incorporated as impurities in crystals. An example of a fluorescence spectrum, figure 3 (*b*), illustrates the splitting of the lines into groups corresponding to different J values. The fluorescence in this solid, CaO(U, Sm) is especially brilliant because of energy transfer from the uranium to the samarium ions. Brauer (1951 a, b) has considered the possible effects of charge compensation on the spectra of rare-earth ions in alkaline-earth oxide lattices. Figure 5 (*a*, *b*) illustrates localized compensation by a vacancy and by a monovalent ion. This has been extended to fluoride lattices

by Stepanov and Feofilov (1956), and figure 5 (*c*, *d*) illustrates compensation by an oxygen ion and by an extra fluorine ion.

One of the peculiar features of the rare-earth salts is that fluorescence has only been reported for ions with from five to ten f-electrons. Promethium with four f-electrons has probably not been closely examined in this respect; but praseodymium and neodymium, and their counterparts in the upper half of the shell, i.e. thulium and erbium, might be expected to fluoresce in the visible or near infrared.† Fluorescence has been observed for all these four elements when incorporated as impurities in various solids, e.g. $CaWO_4$.

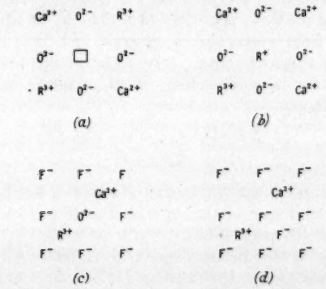


Figure 5. Charge compensation for rare-earth ions in crystal phosphors. Compensation in oxide crystals (*a*) by a vacancy and (*b*) by a monovalent ion, and in fluorite (*c*) by an oxygen ion and (*d*) by a fluorine ion. The normal fluorine ions are shown schematically and really occur in planes above and below the metal ions.

Fluorescence spectra, when they can be observed, enable information to be obtained about the different levels in the lowest multiplet of the ion. Judd (1956) has shown that a partial breakdown of Russell–Saunders coupling occurs even for the ground multiplet, and has used the spectra to obtain values of the spin–orbit splitting constant ξ for the various ions.

4.5. *Trivalent Ions*

$4f^1$: Ce^{3+}, *ground state* $^2F_{5/2}$

The only excited state in the f configuration is $^2F_{7/2}$ at an energy of 2240 cm^{-1} (Lang 1936), which is that due to a spin–orbit splitting constant

$$\xi = 2240 \div 7/2 = 640 \text{ cm}^{-1}.$$

$4f^2$: Pr^{3+}, *ground state* $^3H_4$

This ion has been extensively studied and recent investigations include those of Sayre, Sancier and Freed (1955) on $PrCl_3$, and of Sayre and Freed (1955) on films of PrF_3. Hellwege and Hellwege (1953) have investigated the 'crystal hyperfine structure' ($\sim$0·1 to 1 cm^{-1}) in the spectrum of praseodymium salts.

† It is significant that Dieke and Hall (1957) have found a sharp decrease in the fluorescent lifetimes as one moves away from the half-filled shell.

They attribute this structure to an exchange interaction between praseodymium ions. Brochard and Hellwege (1953) have studied the Zeeman splitting of $Pr_2Mg_3(NO_3)_{12}.24H_2O$.

There are thirteen *LSJ* levels which theory predicts will occur in the following order of increasing energy, as shown in figure 6.

$^3H_4$ $^3H_5$ $^3H_6$ $^3F_2$ $^3F_3$ $^1G_4$ $^3F_4$ $^1D_2$ $^1I_6$ $^3P_0$ $^3P_1$ $^3P_2$ $^1S_0$.

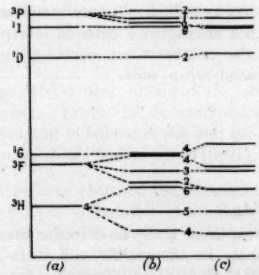


Figure 6. Energy levels for f^3. (*a*) Russell–Saunders levels, (*b*) with first-order spin–orbit interaction, (*c*) with intermediate coupling.

The $^1G_4$ and $^3F_4$ levels are bracketed as they are heavily mixed in the intermediate coupling calculations. These levels have been studied experimentally with the exception of $^3H_5$, $^3H_6$ and $^3F_2$ in the infra-red and $^1S_0$ in the ultra-violet. $^1I_6$ is a doubtful case and has probably not been observed (Jørgensen 1955 c). Figure 6 (*b*) illustrates the splitting of the levels in the Russell–Saunders approximation, whereas figure 5 (*c*) shows the importance of using the matrices of spin–orbit interaction (Spedding 1940) in the intermediate coupling approximation. Judd (1957 b) has calculated the crystal field splittings of the energy levels in $PrCl_3$ using the parameters

$$\xi = 711 \text{ cm}^{-1},$$

$$A_2^0\langle r^2\rangle = \quad 49{\cdot}5 \text{ cm}^{-1} \qquad A_6^0\langle r^6\rangle = -39{\cdot}2 \text{ cm}^{-1},$$

$$A_4^0\langle r^4\rangle = -39{\cdot}7 \text{ cm}^{-1} \qquad A_6^6\langle r^6\rangle = \quad 396{\cdot}7 \text{ cm}^{-1},$$

where $\langle r^n\rangle$ denotes the average value of r^n for a 4f electron. He has found good agreement between theory and experiment.

$4f^3 : Nd^{3+}$, *ground state* $^4I_{9/2}$

There are 41 *LSJ* levels for the ion f^3; but despite the increased complexity good progress has been made with the interpretation of the spectra. Satten (1953) obtained the absorption spectrum of $Nd(BrO_3)_3.9H_2O$ and tentatively identified about 12 of the *LSJ* levels. His interpretation has been discussed further (Jørgensen 1955 d, Satten 1955). The use of intermediate coupling calculations should enable a more positive identification to be made. Ishidzu and Obi (1950)

have calculated the matrices of spin–orbit interaction for f^3, but unfortunately not using the group-theoretical classification of the f-shell (Racah 1949). The matrices have been recalculated using this classification (Loudon, unpublished). Earlier deductions from spectroscopic and specific heat data leading to cubic symmetry for the surroundings of the rare-earth ions in the hydrated sulphates have been shown to be erroneous by Satten and Young's (1955) spectroscopic investigation of $Nd_2(SO_4)_3.8H_2O$ which suggests an orthorhombic or lower symmetry for the ion sites.

The Zeeman effect has been studied in several neodymium salts by Dieke and Heroux (1956), who have not attempted a detailed interpretation of many of the results. With these data the prospects are encouraging for a thorough understanding of the spectra of neodymium salts.

$4f^4$: Pm^{3+}, *ground state* $^5I_4$

Not much information on this ion is available because of the lack of a stable isotope. Parker and Lentz (1950) however find bands of the type expected between 12 000 and 22 000 cm^{-1}.

$4f^5$: Sm^{3+}, *ground state* $^6H_{5/2}$

Gobrecht (1937, 1938) has listed many lines in the absorption and fluorescence spectra of samarium sulphate in the visible and infra-red. By analysing the ground multiplet, Judd (1956) finds $\xi = 1180$ cm^{-1}. Jørgensen (1955 b) predicts ^4M as the lowest quartet state, and hence probably the upper state in fluorescence ; but Elliott, Judd and Runciman (1957) find the means of the lowest quartet multiplets to be, in order of increasing energy, ^4I, ^4M, ^4F and ^4G. When the spin–orbit splitting is taken into account as a first-order perturbation of the Russell–Saunders states, the lowest quartet state is predicted to be $^4F_{3/2}$. The analysis of the spectra of samarium salts is going to be difficult due to the closeness of the lowest quartet states.

The fluorescence of Sm^{3+} has been used by Oshima, Hayakawa, Nagano and Nagusa (1956) to study the phase transition in barium titanate at $-80°$c.

$4f^6$: Eu^{3+}, *ground state* $^7F_0$

Hellwege and Kahle (1951 a, b) have studied the spectra of $EuCl_3.6H_2O$ and of $Eu(BrO_3)_3.9H_2O$, and identified the lowest excited states as $^5D_0$, $^5D_1$ and $^5D_2$. Judd (1955 b) has confirmed these experimental assignments by theoretical study of the energy levels and of the crystal fine structure, and has found $\xi = 1360$ cm^{-1} (Judd 1956).

The fluorescence spectra of europic salts are unique in furnishing three examples of magnetic dipole transitions, namely $^5D_0 \rightarrow {}^7F_1$, $^5D_1 \rightarrow {}^7F_0$, $^5D_1 \rightarrow {}^7F_2$. These transitions have been studied in anhydrous europic chloride (Sayre and Freed 1956 a) and in europic ethylsulphate (Sayre and Freed 1956 b). The magnetic dipole nature of these transitions has also been verified in europic solutions using the technique of wide angle interference (Freed 1942).

$4f^7$: Gd^{3+}, *ground state* $^8S_{7/2}$

This ion has a half-filled shell of 4f electrons, and in consequence exhibits some unusual properties. The states of two highest multiplicities are the ground

state ^{8}S and the excited states ^{6}PDFGHI, all of which have zero spin–orbit and crystal field splitting in the Russell–Saunders approximation. Nutting and Spedding (1937) have examined the ultra-violet absorption spectra of many gadolinium salts, and quite appreciable spin–orbit and crystal field splittings are found which can be attributed to second-order effects. The lowest excited state is probably ^{6}P, although Jørgensen has predicted ^{6}I. The very small splitting of the ground state has been investigated by paramagnetic resonance, and discussed by Hutchison, Judd and Pope (1957). Dieke and Leopold (1957) have studied the absorption, fluorescence and Zeeman spectra of $GdCl_3.6H_2O$ at liquid helium temperatures.†

$4f^8$: Tb^{3+}, *ground state* $^7F_6$

Geisler and Hellwege (1953) have obtained the absorption and fluorescence spectra of terbium bromate. From the fluorescence data Judd (1956) deduces $\xi = 1720$ cm^{-1}.

$4f^9$: Dy^{3+}, *ground state* $^6H_{15/2}$

Dieke and Singh (1956) have obtained the fluorescence and absorption spectra of $DyCl_3.6H_2O$ at 4·2 and 77°K. The interpretation given seems to be invalidated by the theoretical predictions for f^9 made by Elliott, Judd and Runciman (1957). For this ion Judd (1956) finds $\xi = 1920$ cm^{-1}.

$4f^{10}$: Ho^{3+}, *ground state* $^5I_8$

Holmium is of interest because of its correspondence with the naturally missing element promethium which has four f-electrons. From an analysis of the fluorescence spectra of Gobrecht (1938) and Rosenthal (1939), Judd (1956) has deduced $\xi = 2080$ cm^{-1}.

$4f^{11}$: Er^{3+}, *ground state* $^4I_{15/2}$

The absorption spectrum has been studied by Gobrecht (1938). The spin–orbit matrices obtained for f^3 will be applicable to this ion.

$4f^{12}$: Tm^{3+}, *ground state* $^3H_6$

The theoretical interpretation has been discussed by Bethe and Spedding (1937), who found that spin–orbit interactions produce very large corrections, up to 10 000 cm^{-1}, to the Russell–Saunders levels.

$4f^{13}$: Yb^{3+}, *ground state* $^2F_{7/2}$

The absorption spectrum of $YbCl_3.6H_2O$ consists of two sharp lines near 9700Å which are interpreted as transitions to $^2F_{5/2}$. Dieke and Crosswhite (1956) have also examined the Zeeman spectra as a function of crystal orientation.

4.6. *Divalent Ions*

Divalent europium, isoelectronic with Gd^{3+}, has been studied by Freed and Katcoff (1948), Przibram (1949) and by Gruen, Conway and McLaughlin (1956). Freed and Katcoff find the absorption is stronger, by at least a factor of 100, than that of the trivalent ions. This strong absorption is attributed to an allowed

† Cook and Dieke (1957) have now found the $^6P_{3/2}$ group as well as the $^6P_{5/2}$ and $^6P_{7/2}$ groups in gadolinium acetate, the $^6P_{7/2}$ group being lowest and the $^6P_{3/2}$ group being highest in energy.

transition, $4f^7 \rightarrow 4f^6 5d$. Tolstoi, Tkachuk and Tkachuk (1955) find an exponential phosphorescence, $\tau \sim 6 \times 10^{-7}$ sec, which could be consistent with such an assignment. On the other hand, Butement (1948) has studied the absorption spectra of divalent samarium, europium and ytterbium and found not only the broad bands due to allowed transitions, but also narrow weak lines which he correlates with lines of the isoelectronic trivalent rare-earth ions. Feofilov (1956 b) has also studied these divalent ions when present as impurities in fluorite crystals.

§ 5. Heavy Metal Ions

This section deals with ions which have completed sub-shells of electrons in the ground state and are therefore diamagnetic. Such ions in crystals do not usually have a characteristic spectrum in the visible or near ultra-violet, i.e. for wavelengths greater than 2000 Å. However, ions with an ns^2 configuration are readily excited to $nsnp$. Examples of ions with a $6s^2$ configuration are Tl^+, Pb^{2+} and Bi^{3+}, and it is interesting to note that Prener, Hanson and Williams (1953) have used isoelectronic neutral mercury atoms as fluorescent activators. Other ions with an ns^2 configuration are $As^{3+}(4s^2)$, In^+, Sn^{2+} and $Sb^{3+}(5s^2)$, but these ions will not be discussed separately as many of their features will be similar to those of the ions with a $6s^2$ configuration.

(*a*) $6s^2$: Tl^+, *ground state* $^1S_0$

Seitz (1938) interpreted the general features of the spectra of thallium-activated alkali halides in terms of the configuration coordinate diagram. Using this model, Williams (1951) calculated the absorption and emission spectra of KCl(Tl) from the properties and wave functions of the free ions. This calculation, though open to objection on points of detail, has clarified the concept of a luminescent centre. Williams assigns the broad absorption bands at 1960 and 2490 Å to the transitions $^1S_0 \rightarrow {}^1P_1$ and $^1S_0 \rightarrow {}^3P_1$ respectively, and the fluorescence band at 3050 Å to $^3P_1 \rightarrow {}^1S_0$.

Runciman and Steward (1953) have verified by an x-ray investigation of KCl with up to 15% (molecular) of TlCl that the Tl^+ ions enter the lattice substitutionally, and have derived an improved value for the ionic radius. Improved wave functions have been calculated (Douglas, Hartree and Runciman 1955) by the self-consistent field method for Tl^+ in the ground state ($6s^2$) and for the first excited state (6s6p). Unfortunately these wave functions do not include the appreciable exchange and relativistic corrections. These wave functions have been used by Knox and Dexter (1956) to discuss the oscillator strengths in KCl(Tl) and only poor agreement has been obtained with the experimental values. However, Williams, Segall and Johnson (1957) obtain better agreement by including crystalline interactions.

It has been noted that Williams (1953) was mistaken in trying to remove the degeneracy of the $^3P_1$ state by orienting the 6p wave function in the cubic crystal field. It does not now seem clear whether the excited state in the fluorescent transition is the $^3P_0$ state in a field of cubic symmetry or a $^3P_1$ state in an ionic field which has undergone a Jahn–Teller distortion of the type considered by Öpik and Pryce (1957).

The absorption bands of KCl(Tl) have been measured at low temperatures by Johnson and Studer (1951) and by Johnson (1954). The emission and excitation

spectra have been studied by Patterson and Klick (1957) who suggest that the 3050Å and 4750Å emission bands may arise from different centres rather than from two excited states of the same centre as proposed by Johnson and Williams (1952).

(*b*) $6s^2$: Pb^{2+}, *ground state* $^1S_0$

Single crystals of NaCl(Pb) have been studied by Schulman, Ginther and Klick (1950). A vibrational structure has been found in the red fluorescence band in CaO(Pb) at liquid-air temperature (Runciman 1955 a).

(*c*) $6s^2$: Bi^{3+}, *ground state* $^1S_0$

Bismuth-activated oxide phosphors have a well-defined vibrational structure in their fluorescent emission (Ewles and Lee 1953). Runciman (1955 b) has found similar vibrational structure in the absorption band, and has discussed possible localized modes of vibration which are likely to be excited. The absorption and emission bands are situated in the near ultra-violet and blue regions of the spectrum, figure 3 (*c*), and are probably due to transitions between the ground state and an excited ^3P state.

§ 6. Uranium and the Transuranics

The elements from uranium onwards are of special interest because of the part played by 5f electrons in absorption and emission. This section will deal with ions analogous to rare-earth ions and with complexes of the uranyl type.

(*a*) *Uranyl*, $(UO_2)^{2+}$

Dieke and Duncan (1949) have described the known spectroscopic properties of uranyl salts, and in particular the results of isotopic substitution of one or more of the constituents. The absorption spectrum consists of several major series each of which is due to one electronic transition. Each series consists of dozens of lines due to vibrations being excited by the electronic transition. The first series is called the ' fluorescence ' series as it is almost a mirror image of the fluorescence spectrum, allowing that the vibrational frequencies are considerably altered. The symmetrical stretching frequency of the uranyl group which determines the spacing of the line groups in a series is about 880 cm^{-1} in fluorescence and 760 cm^{-1} in absorption, figure 3 (*d*, *e*). The second series in order of increasing energy is the ' magnetic ' series, all the lines of which in uniaxial crystals are split into two by a magnetic field. The ' ultra-violet ' series occurs at still shorter wavelengths.

There is at the moment no detailed theory of the energy spectrum of the uranyl ion, but it seems commonly agreed that in absorption an electron is transferred from an oxygen orbital into the 5f shell which is empty in the ground state. Many of the vibrations can be ascribed to modes characteristic of the uranyl and other groups, e.g. nitrate groups. The width of the lines may be partly due to a small Davydov splitting, as it is probable that the electronic energy is transferred from one uranyl group to another many times before fluorescence occurs. This would account for the sharpness of impurity lines, as in the case of impurities no such mechanism could occur. Further experiments on uranyl salts include a puzzling description by Samoilov (1948) of circular polarization in sodium uranyl

acetate, which is usually regarded as cubic,† and the effect of deuteration on the spectra of uranyl nitrate hydrates (Sevchenko, Vdovenko and Kovaleva 1951).

(*b*) *Hexavalent uranium.*

When many solids are activated with hexavalent uranium they show a green fluorescence at first sight similar to that of uranyl salts. However the spectrum is rather different, figure 3 (*f*), and the idea of localized charge compensation suggests that a UO_4 group is present in $CaF_2(U, O)$, figure 7 (*a*), and a UO_6 group in CaO(U) (Runciman 1955 a). The concept of a fluorescent UO_6 group is strengthened by the existence of fluorescent alkaline earth uranates such as Ba_2CaUO_6 (Steward and Runciman 1953).

When NaF(U) is prepared under oxidizing conditions, as is necessary for it to be luminescent, then a UO_6 group, figure 7 (*b*), can be used to account for the vibrational frequencies in the fluorescence spectrum and their change with ^{18}O substitution (Runciman 1956 b). The fluorescence spectrum of NaF(U, O) changes considerably on cooling from 77 to 4·2°K (figure 8 (*a*, *b*)) indicating that the transition probability from one of the upper excited states is greater than from the lowest excited state. The absorption spectrum obtained by the reflection method, figure 8 (*c*), shows this upper excited state now in absorption.

(*c*) *Lower valence states of uranium.*‡

Uranium can readily occur in the trivalent and tetravalent states. Jørgensen (1955 a) attributes the U^{4+} spectrum to an f^2 configuration. Alexanian (1957) has described the absorption spectrum of some tetravalent uranium compounds. Galkin and Feofilov (1957) have found an infra-red luminescence of U^{3+}, and it will be interesting to see if results on U^{3+} are analogous to the rare earth ion ND^{3+}.

(*d*) *Neptunium.*‡

Freed and Leitz (1949) briefly describe the absorption spectrum of $NpCl_4$, for which the electronic configuration may be f^3.

Neptunyl salts have also been studied, and there are two bands with vibrational structure at 18 000 and 21 200 cm^{-1}. Eisenstein and Pryce (1955) suppose that the axial field in the $(NpO_2)^{2+}$ ion dominates and that the absorption spectrum is due to the one 5f electron in a non-bonding state. The ground state is $\phi_{\pm 5/2}$ and the above absorption bands are attributed to transitions to $\sigma_{\pm 5/2}$ and $\sigma_{\pm 3/2}$, where σ and ϕ stand for states with $l_z = 2, 3$ respectively.

(*e*) *Plutonium.*

Freed and Leitz (1949) have obtained the absorption spectrum of $PuCl_3$, and Cunningham, Gruen, Conway and McLaughlin (1956) have discovered and described the fluorescence of Pu^{3+} as an activator in $LaCl_3$. No analysis of these spectra is available, but the ion is probably analogous to Sm^{3+}, which has an f^5 configuration. Leontovich (1957) has obtained the absorption spectrum of tetravalent plutonium in several salts, e.g. $Pu(SO_4)_2 . 4H_2O$.

† Nye (1957) includes a most helpful account of optical activity in crystals. The cubic classes O and T show optical activity, and sodium acetate probably belongs to the latter class.

‡ Conway, Wallmann, Cunningham and Shalimoff (1957) have found fluorescence spectra for trivalent uranium, neptunium and curium in $LaCl_3$, and an ultra-violet absorption spectrum for curium in $LaCl_3$. The uranium fluorescence is in the range 5 400–7 000 Å and is obviously different in origin from that found by Galkin and Feofilov (1957).

Eisenstein and Pryce (1956) have extended their calculations on uranyl-type ions to plutonyl, and attribute many bands to transitions within the f^2 configuration. By comparison with the uranyl ion, Jørgensen (1957) thinks some of these bands may be due to a process in which an electron is transferred into the 5f shell.

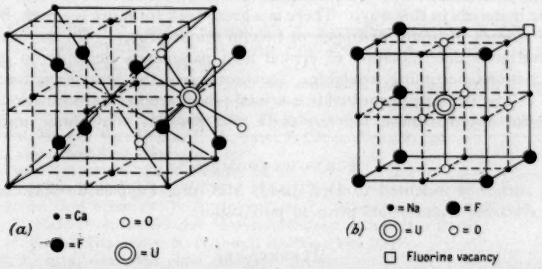


Figure 7. Charge compensation in uranium-activated fluorides. (*a*) CaF_2(U, O), (*b*) NaF(U, O).

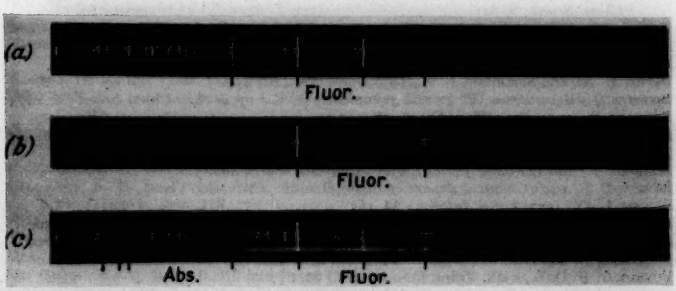


Figure 8. Crystal spectra obtained using a large glass Littrow spectrograph. For calibration purposes an iron-arc spectrum has been superimposed on the central portion of each spectrum. (*a*) Fluorescence spectrum of NaF(U, O) at 77°K. (*b*) Fluorescence spectrum of NaF(U, O) at 4·2°K. (*c*) Absorption spectrum of NaF(U, O) at 77°K.

(*f*) *Americium.*

$AmCl_3$ has a sharp line spectrum (Freed and Leitz 1949), and appears closely analogous to Eu^{3+}, f^6. The fluorescence of Am^{3+} in $LaCl_3$ has been described by Gruen, Conway, McLaughlin and Cunningham (1956), but no detailed analysis is yet available for this ion or for the americyl ion.

(*g*) *Curium.*

Trivalent curium may be expected to be similar to trivalent gadolinium f^7, and to show absorption and fluorescence in the ultra-violet. Curium does not form a stable complex of the uranyl type.

§ 7. Conclusion

The interpretation of fluorescence and absorption spectra is leading to rapid progress in our knowledge of the energy levels of ions in solids. This will extend the possible applications of spectroscopy to solid-state physics. In particular, it should be possible to study the phase changes in some ferroelectric and antiferromagnetic materials in this way. There is a great need for more research, including Zeeman studies, at liquid hydrogen or helium temperatures. On the other hand, the quantitative determination of crystal field parameters depends on improved wave functions becoming available. In most cases the wave functions of the free ion will be strongly perturbed in a crystal ; but free ion wave functions should be a useful approximation for rare-earth and possibly transuranic ions.

Acknowledgments

The author is indebted to Drs. Judd, McClure, Pappalardo and Satten for making available manuscripts prior to publication.

References

Alexanian, V. T., 1957, *Doklady Akad. Nauk. S.S.S.R.*, **115**, 333.
Ballhausen, C. J., 1954, *K. Danske Vidensk. Selsk. Mat.-fys. Medd.*, **29**, No. 4 ; 1955, *Acta Chem. Scand.*, **9**, 821.
Ballhausen, C. J., and Jørgensen, C. K., 1955, *K. Danske Vidensk. Selsk. Mat.-fys. Medd.*, **29**, No. 14.
Bethe, H., 1929, *Ann. Phys., Lpz.*, **3**, 133 ; 1930, *Z. Phys.*, **60**, 218.
Bethe, H. A., and Spedding, F. H., 1937, *Phys. Rev.*, **52**, 454.
Bjerrum, J., Ballhausen, C. J., and Jørgensen, C. K., 1954, *Acta Chem. Scand.*, **8**, 1275.
Bleaney, B., 1955, *Proc. Phys. Soc.* A, **68**, 937.
Bleaney, B., and Stevens, K. W. H., 1953, *Rep. Progr. Phys.*, **16**, 108 (London : Physical Society).
Bouriau, Y., and Lenoble, J., 1957, *C.R. Acad. Sci., Paris*, **245**, 511.
Bowen, E. J., 1946, *Chemical Aspects of Light* (Oxford : Clarendon Press).
Bowen, L. O., 1957, *J. Sci. Instrum.*, **34**, 265.
Bowers, K. D., and Owen, J., 1955, *Rep. Progr. Phys.*, **18**, 304 (London : Physical Society).
Brauer, P., 1951 a, *Z. Naturf.*, **6a**, 561 ; 1951 b, *Ibid.*, **6a**, 562.
Brochard, J., and Hellwege, K. H., 1953, *Z. Phys.*, **135**, 620.
Butement, F. D. S., 1948, *Trans. Faraday Soc.*, **44**, 617.
Condon, E. U., and Shortley, G. H., 1935, *The Theory of Atomic Spectra* (Cambridge : University Press).
Conway, J. G., Wallmann, J. C., Cunningham, B. B., and Shalimoff, G. V., 1957, *J. Chem. Phys.*, **27**, 1416.
Cook, S. P., and Dieke, G. H., 1957, *J. Chem. Phys.*, **27**, 1213.
Couture, L., 1957, *C.R. Acad. Sci., Paris*, **245**, 515.
Couture, L., Jacquinot, P., and Tsujikawa, I., 1956, *Brit. J. Appl. Phys.*, **7**, 425.
Cunningham, B. B., Gruen, D. M., Conway, J. C., and McLaughlin, R. D., 1956, *J. Chem. Phys.*, **24**, 1275.
Deutschbein, O., 1939, *Ann. Phys., Lpz.*, **36**, 183.
Deutschbein, O., Joos, G., and Teltow, J., 1942, *Naturwissenschaften*, **30**, 228.
Dexter, D. L., Klick, C. C., and Russell, G. A., 1955, *Phys. Rev.*, **100**, 603.
Dieke, G. H., and Crosswhite, H. M., 1956, *J. Opt. Soc. Amer.*, **46**, 885.
Dieke, G. H., and Duncan, A. B. F., 1949, *Spectroscopic Properties of Uranium Compounds*, National Nuclear Energy Series, III, 2 (New York : McGraw-Hill).
Dieke, G. H., and Hall, L. A., 1957, *J. Chem. Phys.*, **27**, 465.
Dieke, G. H., and Heroux, L., 1956, *Phys. Rev.*, **103**, 1227.

DIEKE, G. H., and LEOPOLD, L., 1957, *J. Opt. Soc. Amer.*, **47**, 944.
DIEKE, G. H., and SINGH, S., 1956, *J. Opt. Soc. Amer.*, **46**, 495.
DOUGLAS, A. S., HARTREE, D. R., and RUNCIMAN, W. A., 1955, *Proc. Camb. Phil. Soc.*, **51**, 486.
DREISCH, T., and KALLSCHEUER, O., 1944, *Z. Elektrochem.*, **50**, 224.
EISENSTEIN, J. C., and PRYCE, M. H. L., 1955, *Proc. Roy. Soc.* A, **229**, 20; 1956, *Ibid.*, **238**, 31.
ELLIOTT, J. P., JUDD, B. R., and RUNCIMAN, W. A., 1957, *Proc. Roy. Soc.* A, **240**, 509.
ELLIOTT, R. J., and STEVENS, K. W. H., 1952, *Proc. Roy. Soc.* A, **215**, 437; 1953 a, *Ibid.*, **218**, 553; 1953 b, *Ibid.*, **219**, 387.
ELLIS, J. W., and KAPLAN, J., 1935, *J. Opt. Soc. Amer.*, **25**, 357; 1937, *Ibid.*, **27**, 94.
EWLES, J., and LEE, N., 1953, *J. Electrochem. Soc.*, **100**, 392, 399, 402.
FEOFILOV, P. P., 1956 a, *J. Phys. Radium* (8), **17**, 656; 1956 b, *Optika I Spektroskopiya*, **1**, 992.
FINKELSTEIN, R., and VAN VLECK, J. H., 1940, *J. Chem. Phys.*, **8**, 790.
FREED, S., 1942, *Rev. Mod. Phys.*, **14**, 105.
FREED, S., and KATCOFF, S., 1948, *Physica*, **14**, 17.
FREED, S., and LEITZ, F. J., 1949, *J. Chem. Phys.*, **17**, 540.
GALKIN, L. N., and FEOFILOV, P. P., 1957, *Doklady Akad. Nauk. S.S.S.R.*, **114**, 745.
GARLICK, G. F. J., 1949, *Luminescent Materials* (Oxford: Clarendon Press).
GEISLER, H. F., and HELLWEGE, K. H., 1953, *Z. Phys.*, **136**, 293.
GIELESSEN, J., 1935, *Ann. Phys., Lpz.*, **22**, 537.
GOBRECHT, H., 1937, *Ann. Phys., Lpz.*, **28**, 673; 1938, *Ibid.*, **31**, 755.
GRECHUSHNIKOV, B. N., and FEOFILOV, P. P., 1955, *J. Exp. Theor. Phys.*, **29**, 384 (translated in *Soviet Physics JETP*, 1956, **2**, 330).
GRUEN, D. M., CONWAY, J. G., and MCLAUGHLIN, R. D., 1956, *J. Chem. Phys.*, **25**, 1102.
GRUEN, D. M., CONWAY, J. G., MCLAUGHLIN, R. D., and CUNNINGHAM, B. B., 1956, *J. Chem. Phys.*, **24**, 1115.
HAMEKA, H. F., and VLAM, C. C., 1953, *Physica*, **19**, 943.
HARTMANN, H., and SCHLÄFER, H. L., 1951 a, *Z. Phys. Chem.*, **197**, 115; 1951 b, *Z. Naturf.*, **6a**, 760.
HELLWEGE, A. M., 1955, LANDOLT-BÖRNSTEIN, *Zahlenwerte and Funktionen*, 6th edn, Vol. 1, Part 4, 893 (Berlin: Springer).
HELLWEGE, A. M., and HELLWEGE, K. H., 1953, *Z. Phys.*, **135**, 615.
HELLWEGE, K. H., 1948, *Ann. Phys., Lpz.*, **4**, 95, 127, 136, 143, 150; 1950 a, *Z. Phys.*, **127**, 513; 1950 b, *Ibid.*, **128**, 172.
HELLWEGE, K. H., and KAHLE, H. G., 1951 a, *Z. Phys.*, **129**, 62; 1951 b, *Ibid.*, **129**, 85.
HOLMES, O. G., and MCCLURE, S., 1957, *J. Chem. Phys.*, **26**, 1686.
HUSH, N. S., and PRYCE, M. H. L., 1957, *J. Chem. Phys.*, **26**, 143.
HUTCHISON, C. A., JUDD, B. R., and POPE, D. F. D., 1957, *Proc. Phys. Soc.* B, **70**, 514.
ISHIDZU, T., and OBI, S., 1950, *J. Phys. Soc. Japan*, **5**, 145.
JOHNSON, P. D., 1954, *J. Chem. Phys.*, **22**, 1143.
JOHNSON, P. D., and STUDER, F. J., 1951, *Phys. Rev.*, **82**, 976.
JOHNSON, P. D., and WILLIAMS, F. E., 1952, *J. Chem. Phys.*, **20**, 124.
JØRGENSEN, C. K., 1955 a, *K. Danske Vidensk. Selsk. Mat.-fys. Medd.*, **29**, No. 7; 1955 b, *Ibid.*, **29**, No. 11; 1955 c, *Acta Chem. Scand.*, **9**, 540; 1955 d, *J. Chem. Phys.*, **23**, 399; 1957, *Acta Chem. Scand.*, **11**, 166.
JUDD, B. R., 1955 a, *Proc. Roy. Soc.* A, **227**, 552; 1955 b, *Ibid.*, **228**, 120; 1955 c, *Ibid.*, **232**, 458; 1956, *Proc. Phys. Soc.* A, **69**, 157; 1957 a, *Proc. Roy. Soc.* A, **241**, 122; 1957 b, *Ibid.*, **241**, 414.
KIRK, R. D., 1954, *J. Electrochem. Soc.*, **101**, 461.
KLICK, C. C., and SCHULMAN, J. H., 1952, *J. Opt. Soc. Amer.*, **42**, 910.
KNOX, R. S., and DEXTER, D. L., 1956, *Phys. Rev.*, **104**, 1245.
KRISHNAN, R. S., 1947, *Proc. Indian Acad. Sci.*, **26A**, 450.
KRÖGER, F. A., 1948, *Some Aspects of the Luminescence of Solids* (Amsterdam: Elsevier).
LANG, R. J., 1936, *Canad. J. Res.*, **14**, 127.
LEONTOVICH, A. M., 1957, *Optika I Spektroskopiya*, **2**, 695.

LIEHR, A. D., and BALLHAUSEN, C. J., 1957, *Phys. Rev.*, **106**, 1161.
LOW, W., 1957a, *Z. Phys. Chem.*, **13**, 107 ; 1957 b, *Phys. Rev.*, **105**, 807.
MANI, A., 1942, *Proc. Indian Acad. Sci.*, **15A**, 52.
MOFFITT, W., and BALLHAUSEN, C. J., 1956, *Annual Review of Physical Chemistry*, **7**, 107 (Palo Alto : Annual Reviews).
NUTTING, G. C., and SPEDDING, F. H., 1937, *J. Chem. Phys.*, **5**, 33.
NYE, J. F., 1957, *Physical Properties of Crystals* (Oxford : Clarendon Press).
ÖPIK, U., and PRYCE, M. H. L., 1957, *Proc. Roy. Soc.* A, **238**, 425.
ORGEL, L. E., 1952, *J. Chem. Soc.*, Pt. IV, 4756 ; 1955 a, *J. Chem. Phys.*, **23**, 1004 ; 1955 b, *Ibid.*, **23**, 1819 ; 1955 c, *Ibid.*, **23**, 1824 ; 1955 d, *Ibid.*, **23**, 1958.
ORGEL, L. E., and DUNITZ, J. D., 1957, *Nature, Lond.*, **179**, 462.
OSHIMA, K., HAYAKAWA, S., NAGANO, H., and NAGUSA, M., 1956, *J. Chem. Phys.*, **24**, 903.
OWEN, J., 1955, *Proc. Roy. Soc.* A, **227**, 183.
PAPPALARDO, R., 1957 a, *Nuovo Cim.*, **6**, 392 ; 1957 b, *Phil. Mag.* (8) **2**, 1397.
PARKER, G. W., and LENTZ, P. M., 1950, *J. Amer. Chem. Soc.*, **72**, 2834.
PATTERSON, D. A., and KLICK, C. C., 1957, *Phys. Rev.*, **105**, 401.
PRENER, J. S., 1953, *J. Chem. Phys.*, **21**, 160.
PRENER, J. S., HANSON, R. E., and WILLIAMS, F. E., 1953, *J. Chem. Phys.*, **21**, 759.
PRINGSHEIM, P., 1949, *Fluorescence and Phosphorescence* (New York : Interscience Publishers).
PRZIBRAM, K., 1949, *Nature, Lond.*, **163**, 989 ; 1956, *Irradiation Colours and Luminescence* (translated by J. E. Caffyn) (London : Pergamon Press).
RACAH, G., 1949, *Phys. Rev.*, **76**, 1352.
ROSENTHAL, G., 1939, *Phys. Z.*, **40**, 508.
RUNCIMAN, W. A., 1955 a, *Brit. J. Appl. Phys.*, **6**, Suppl. No. 4, 78 ; 1955 b, *Proc. Phys. Soc.* A, **68**, 647 ; 1956 a, *Phil. Mag.* (8), **1**, 1075 ; 1956 b, *Proc. Roy. Soc.* A, **237**, 39.
RUNCIMAN, W. A., and STEWARD, E. G., 1953, *Proc. Phys. Soc.* A, **66**, 484.
SAMOILOV, B. N., 1948, *J. Exp. Theor. Phys.*, **18**, 1030.
SATTEN, R. A., 1953, *J. Chem. Phys.*, **21**, 637 ; 1955, *Ibid.*, **23**, 400 ; 1957, *Ibid.*, **27**, 286.
SATTEN, R. A., and YOUNG, D. J., 1955, *J. Chem. Phys.*, **23**, 404.
SAYRE, E. V., and FREED, S., 1955, *J. Chem. Phys.*, **23**, 2066 ; 1956 a, *Ibid.*, **24**, 1211 ; 1956 b, *Ibid.*, **24**, 1213.
SAYRE, E. V., SANCIER, K. M., and FREED, S., 1955, *J. Chem. Phys.*, **23**, 2060.
SCHULMAN, J. H., GINTHER, R. J., and KLICK, C. C., 1950, *J. Opt. Soc. Amer.*, **40**, 854.
SEITZ, F., 1938, *J. Chem. Phys.*, **6**, 150.
SEVCHENKO, A. N., VDOVENKO, V. M., and KOVALEVA, T. V., 1951, *J. Exp. Theor. Phys.*, **21**, 204.
SPEDDING, F. H., 1940, *Phys. Rev.*, **58**, 255.
SPEDDING, F. H., and NUTTING, G. C., 1934, *J. Chem. Phys.*, **2**, 421 ; 1935, *Ibid.*, **3**, 369.
STEPANOV, I. V., and FEOFILOV, P. P., 1956, *Doklady Akad. Nauk S.S.S.R.*, **108**, 615 (translated in *Soviet Physics Doklady*, **1**, 350).
STEVENS, K. W. H., 1952, *Proc. Phys. Soc.* A, **65**, 209.
STEWARD, E. G., and RUNCIMAN, W. A., 1953, *Nature, Lond.*, **172**, 75.
TANABE, Y., and SUGANO, S., 1954 a, *J. Phys. Soc. Japan*, **9**, 753 ; 1954 b, *Ibid.*, **9**, 766 ; 1957, *Ibid.*, **12**, 556.
TELTOW, J., 1938, *Z. Phys. Chem.* B, **40**, 397 ; 1939, *Ibid.*, **43**, 198.
TOLSTOI, N. A., TKACHUK, A. M., and TKACHUK, N. N., 1955, *J. Exp. Theor. Phys.* **29**, 386 (translated in *Soviet Physics JETP*, 1956, **2**, 331).
VAN VLECK, J. H., 1937, *J. Phys. Chem.*, **41**, 67 ; 1940, *J. Chem. Phys.*, **8**, 787.
WILLIAMS, F. E., 1951, *J. Chem. Phys.*, **19**, 457 ; 1953, *J. Phys. Chem.*, **57**, 780.
WILLIAMS, F. E., SEGALL, B., and JOHNSON, P. D., 1957, *Phys. Rev.*, **108**, 46.
WOLFSBERG, M., and HELMHOLZ, L., 1952, *J. Chem. Phys.*, **20**, 837.
YAMADA, S., 1951, *J. Amer. Chem. Soc.*, **73**, 1182.
YOST, D. M., RUSSELL, H., and GARNER, C. S., 1947, *The Rare-earth Elements and their Compounds* (New York : Wiley).

ELECTROMAGNETIC EXCITATION OF NUCLEI BY NUCLEAR PROJECTILES AND ELECTRONS

By R. HUBY
University of Liverpool

CONTENTS

	PAGE
List of symbols	60
§ 1. Introduction	62
PART I: THEORY	
§ 2. The interaction	63
§ 3. Approximate methods	66
3.1. Semi-classical method	66
3.2. Born approximation	68
3.3. Method of Coulomb-distorted waves	69
§ 4. Order-of-magnitude estimates	69
4.1. Formulae for total cross section	69
4.2. Discussion of total cross section	71
4.3. Angular distribution of inelastically scattered projectiles	73
§ 5. Excitation by nuclear projectiles	76
5.1. Semi-classical method	76
5.2. Quantal calculation	78
5.3. Limitations of present theory	81
5.4. J.B.W.K. method: bridge between semi-classical and quantal theories	83
5.5. Angular distribution of de-excitation γ-rays	84
§ 6. Excitation by electrons	87
6.1. Electrostatic interaction in Born approximation	87
6.2. Sum of cross sections over states	89
6.3. Complete electromagnetic interaction	90
6.4. Further refinements	91
PART II: EXPERIMENT	
§ 7. Excitation by nuclear projectiles	92
7.1. Experimental methods	92
7.2. Optimum experimental conditions	93
7.3. Observation of various multipoles	95
7.4. Nuclear data obtained from excitation experiments	98
§ 8. Excitation by electrons	100
8.1. Experimental method	100
8.2. Optimum experimental conditions	101
8.3. Observation of various multipoles	102
8.4. Nuclear data obtained from excitation experiments	103
§ 9. Comparison of electrons and nuclear projectiles	106
Appendix: Magnetic excitation	107
Acknowledgment	109
References	109

Abstract. The principles of the excitation of nuclei by their electromagnetic interaction with bombarding particles are developed in §§ 1–4 in a common way for electrons and nuclear projectiles. Thereafter these two types of projectile are dealt with separately, first as regards theory and then the experimental method.

List of Symbols

The equation or section where the symbol is introduced is shown in brackets.

a. $\frac{1}{2}$ classical closest distance of approach. (3.9)
$a_k^{E\lambda}$. Eλ decay, γ-angular distribution coefficient. (5.25)
A_1, A_2. Mass number of projectile and target nucleus respectively. (§ 2)
A_k. γ–γ correlation coefficient. (5.29)
$\mathbf{A}(\mathbf{r}_p)$. Vector potential due to nucleus. (2.3)
b_{if}. Semi-classical excitation amplitude. (3.1)
$B(\lambda) = B(E\lambda)$ or $B(M\lambda)$ according to circumstances. (4.16)
$B(E\lambda), B(M\lambda)$. Reduced radiative probability, electric or magnetic respectively. (4.6), (A4)
$B(C\lambda, K)$. Reduced probability for electrostatic excitation by electrons. (6.5)
$B(E\lambda, K)$. As last, but for Eλ excitation due to $V_A^{(1)}$. (6.21)
$B(M\lambda, K)$. As last, but for magnetic excitation. (A7)
$C_{k0}(\lambda\lambda)$. Tensor radiation parameter. (5.25)
E_B. Coulomb barrier. (4.17)
E_i, E_f. Initial or final nuclear energy level. (3.2)
Eλ. λ-order electric multipole. (§ 2)
ΔE. Excitation energy. (§ 2)
ΔE_{max}. Maximum ΔE excitable by projectile beam. (§ 7.2)
$f_{E\lambda}(\xi), f_{M\lambda}(\xi)$. Dimensionless function in total cross section, semi-classical, electric or magnetic respectively, (nuclear projectiles). (5.3), (A4)
$df_{E\lambda}(\theta, \xi)$, $df_{M\lambda}(\theta, \xi)$. As last, but for differential cross section. (5.6), (App.)
$f_{E\lambda}(\eta_i, \xi)$, $f_{M\lambda}(\eta_i, \xi)$. Dimensionless function in total cross section, quantal, electric or magnetic respectively. (5.10), (App.)
$df_{E\lambda}(\theta, \eta_i, \xi)$, $df_{M\lambda}(\theta, \eta_i, \xi)$. As last, but for differential cross section. (5.8), (App.)
$F(I_i \to I_f, K)$. Form factor, electrostatic, for electrons. (6.4)
$|F_{E\lambda}(I_i \to I_f, K)|^2$, $|F_{M\lambda}(I_i \to I_f, K)|^2$. Multipolar form factors (square), electric or magnetic respectively. (6.5), (A7)
$|F_{E\lambda(A)}(I_i \to I_f, K)|^2$. E$\lambda$ form factor due to $V_A^{(1)}$ (square). (6.21)
$|F|_{el}^2$. Elastic form factor (square). (6.16)
$|F|_{Inel}^2$. Summed inelastic form factor (square). (6.17)
$F_l(kr)$. Radial Coulomb wave function. (5.11)
I_i, I_f. Initial or final nuclear spin. (4.6)
$I_{\lambda\mu}(\theta, \xi)$. Dimensionless orbital excitation integral. (5.2)
$j_\lambda(x)$. Spherical Bessel function. (6.6)
$\mathbf{j}_n$. Nuclear current density. (2.3)
k (suffix). Label of k'th proton in nucleus (2.1); also tensor order in angular distribution. (5.24)
k. Wave number of projectile: mean of k_i, k_f. (3.10)
$\mathbf{k}_i, \mathbf{k}_f$. Initial or final projectile wave number. (3.13)
$\mathbf{K} = \mathbf{k}_i - \mathbf{k}_f$. (4.22)
l. Projectile's orbital angular momentum: mean of l_i, l_f. (5.19)
l_i, l_f. Initial or final projectile's orbital angular momentum. (5.11)
$\mathbf{L}$. Rotational operator. (A3)
m_0. Reduced mass. (§ 3.1)

m_p. Projectile mass. (3.9)
M. Proton mass. (2.5)
M_i, M_f. Magnetic quantum number of initial or final level. (4.6)
$\mathrm{M}\lambda$. λ-order magnetic multipole. (§ 2)
$M_{fi}{}^{-\lambda-1}$. Radial matrix element. (5.11)
$\mathfrak{M}(\mathrm{M}\lambda, \mu)$. Radiative magnetic multipole operator. (A2)
$\mathfrak{M}(\mathrm{M}\lambda, \mu, \mathrm{K})$. Magnetic multipole operator for electron excitation. (A8)
p. Impact parameter. (3.3)
$p_{\min}$. Lower cut-off value of p. (§ 3.1)
$p_{\max}$. Upper cut-off value of p. (4.9)
$r_{\min}$. Distance of closest approach on a particular orbit. (§ 4.3)
$r_{p,\max}$. Maximum distance for interaction. (2.8)
$\boldsymbol{r}$. Point in nucleus. (2.3)
$\boldsymbol{r}_k$. Coordinate of kth proton. (2.1)
$\boldsymbol{r}_p$. Coordinate of projectile. (2.1)
$R(\eta_i, \xi)$. Ratio of quantal to semi-classical cross section. (§ 5.2)
R_0. Nuclear radius. (4.15)
$S_{\mathrm{E}\lambda,\mu}$. Orbital excitation integral. (4.2)
T. Projectile's kinetic energy, far from nucleus (mean). (3.6)
$T_{\min}$. Least value of T required to excite transition. (7.1)
T_λ. Rate of λ-order radiative decay. (4.16)
v. Projectile's velocity far from nucleus : mean of v_i, v_f. (3.8)
v_i, v_f. Initial or final projectile velocity. (3.13)
$\mathbf{v}_k$. Velocity of kth nucleon. (§ 2)
$\mathbf{v}_p$. Velocity of projectile (instantaneous). (2.2)
V. Electrostatic interaction. (2.1)
$V' = V - V_0$. (3.5)
$V_A{}^{(1)}$. Interaction due to convection current of projectile. (2.2)
$V_A{}^{(2)}$. Interaction due to magnetic moment of projectile. (2.5)
$V_M{}^{(1)}$. Magnetic part of $V_A{}^{(1)}$. (A1)
V_0. Point charge electrostatic interaction. (3.4)
$W(\theta)$. γ–γ correlation function. (5.29)
$W(\theta\gamma)$. Decay γ-angular distribution. (5.30)
Z_1, Z_2. Atomic numbers of projectile and nucleus respectively. (§ 2.1)
η. Coulomb scattering parameter : mean of η_i, η_f. (3.8)
η_i, η_f. Initial and final values of η. (5.5)
θ. Angle of scattering. (3.11)
θ_k. Polar coordinate of kth proton. (2.9)
θ_p. Polar coordinate of projectile. (2.9)
θ_γ. Angle of emission of γ-ray. (5.30)
κ (suffix). Labels tensor component in angular distribution. (5.24)
κ. Photon wave number : $\Delta E/\hbar c$. (2.7)
λ. Multipole order. (2.9)
$\lambda\!\!\!^{-}$. Wavelength of projectile : $1/k$. (3.10)
$(2\lambda+1)!! = 1.3.5.\ldots.(2\lambda+1)$. (4.14)
μ. Magnetic quantum number going with λ. (2.9)
$\boldsymbol{\mu}_p$. Magnetic moment of projectile. (2.5)

ξ. Adiabatic parameter for process. (4.12) or (5.5)
$\xi(p)$. Adiabatic parameter for one orbit. (4.3)
$\rho_c(\mathbf{r}_1, \mathbf{r}_2)$. Correlation density in ground state. (6.19)
$\rho_1(\mathbf{r})$. One-proton density in ground state. (6.18)
$\rho_{k,\kappa}$. Statistical tensor of excited state. (5.24)
$\rho_{M_f,M_f'}$. Density matrix of excited state. (5.23)
$\rho_\lambda(I_i \to I_f, r)$. λ-order radial transition density. (6.10)
$\sigma(\theta, \phi)$. Differential cross section. (3.11)
$\sigma_{E\lambda}, \sigma_{M\lambda}$. λ-order total cross section, electric or magnetic respectively. (4.8), (A4)
$\sigma_{E\lambda}(\theta)$, $\sigma_{M\lambda}(\theta)$. λ-order differential cross section, electric or magnetic respectively. (4.20). (App.)
σ_R. Rutherford cross section. (3.12)
ϕ. Azimuthal scattering angle. (3.3)
$\phi_i, \phi_f, \tilde{\phi}_f$. Initial, final and modified projectile Coulomb wave functions. (3.15)
ϕ_k. Polar coordinate of kth proton. (2.9)
ϕ_p. Polar coordinate of projectile. (2.9)
ψ_i, ψ_f. Initial and final nuclear wave functions. (3.1)
ω. Transition frequency $\Delta E/\hbar$. (3.2)

§ 1. Introduction

This report deals with the collision of charged particles with nuclei in which the nucleus is raised to an excited state and the projectile inelastically scattered as a result of the electromagnetic interaction between the charges and, possibly, magnetic moments. The projectiles which have been used hitherto are the light nuclear particles proton, deuteron and alpha particle, the heavier nucleus ^{14}N (or corresponding ions) and the electron. The process is commonly referred to as 'Coulomb excitation', and indeed in all important cases the operative interaction reduces to the simple electrostatic Coulomb potential between the charges of the projectile and the protons of the nucleus. A more sophisticated treatment shows the process as the virtual emission of a photon by the projectile and its absorption by the target (or vice versa). This description indicates that an essential role is played by the electromagnetic transition probability $B(\lambda)$ of the nucleus for emission or absorption of a γ-ray in the transition concerned. Consequently the process is capable of yielding values of these probabilities, and information as to the multipole character λ of the transitions. Magnetic transitions Mλ as well as electric Eλ can in principle be excited, but so much more weakly that they have never been observed, and discussion of magnetic effects is relegated to an appendix in this report.

Excitation of a nuclear level by electrons can result only from the electromagnetic interaction, but when nuclear projectiles are used, the same result can be brought about also by nuclear forces. To minimize this complication, the energy of nuclear projectiles is normally kept below the Coulomb barrier; the long range of Coulomb forces permits Coulomb excitation to occur when the projectile is outside the nucleus. The most favourable energies for electrons, on the other hand, are around 100 MeV or more.

Coulomb excitation by nuclear projectiles has received a great deal of attention recently from theoreticians and experimentalists, following the earlier studies of Weisskopf (1938), Mullin and Guth (1951), Huby and Newns (1951) and Ter-Martirosyan (1952). Experimentally over 200 transitions, almost all electric quadrupole, have been investigated. The method is most suitable for exciting low-lying levels with large E2 transition moments. Many transition probabilities B(E2) have been determined, and these are often too large, and the corresponding life-times too small ($\lesssim 10^{-9}$ sec), to be readily accessible to other experimental methods. The nuclei cover all ranges of mass but the lightest, and the data obtained have proved of great value for the substantiation of the collective model of nuclei (see the article by Brink in the present volume). A very comprehensive review of Coulomb excitation by nuclear projectiles has been given by Alder, Bohr, Huus, Mottelson and Winther (1956), and this has been treated as an authoritative reference for the present report. An excellent review of the more experimental aspects of the subject has been made by Heydenburg and Temmer (1956 a).

The inelastic scattering of electrons is much less developed. The theory has only been worked out in Born approximation, which is only vaid for light nuclei. The electrons penetrate freely through the nucleus, giving rise to 'probing' effects like those more familiar in the elastic scattering of electrons, and thus yielding more detailed information about the nuclear wave functions than nuclear projectiles do. Also electrons alone can excite 0–0 transitions. To date (mid 1957), about a dozen levels of several MeV excitation have been excited in light-to-medium weight nuclei with 200 MeV electrons at Stanford. Rather tentative transition probabilities and multipolarities have been deduced, and a number of theoretical papers relating these results to nuclear models have been written.

Coulomb excitation by μ-mesons would be an interesting possibility, but it is discouraged by evidence that the cross section for scattering of high-energy (> 200 MeV) μ-mesons by heavy nuclei at large angles is many times larger than could theoretically result from electromagnetic excitation (see for example Fowler 1956).

The present report deals with the principles of the electromagnetic excitation process rather than the particular data which have been obtained from it.

PART I. THEORY

§ 2. The Interaction

We consider the electromagnetic interaction between a projectile of charge $Z_1 e$, which is regarded as a point charge, and a nucleus of charge $Z_2 e$. In the simplest electrostatic approximation this is the Coulomb interaction

$$V = Z_1 e^2 \sum_k \frac{1}{|\mathbf{r}_p - \mathbf{r}_k|}, \qquad \ldots\ldots(2.1)$$

where $\mathbf{r}_p$ is the coordinate of the projectile, and the sum extends over the protons k of the nucleus. To obtain also magnetic effects, we may write down a further interaction $V_A^{(1)}$ of the convection current of the projectile with the vector potential due to the nucleus :

$$V_A^{(1)} = -\frac{Z_1 e \mathbf{v}_p . \mathbf{A}(\mathbf{r}_p)}{c}, \qquad \ldots\ldots(2.2)$$

where $\mathbf{v}_p$ is the velocity of the projectile, and for the vector potential $\mathbf{A}(\mathbf{r}_p)$ we use provisionally the naïve classical formula

$$\mathbf{A}(\mathbf{r}_p) = \frac{1}{c}\int \frac{\mathbf{j}_n(\mathbf{r})\, d\tau}{|\mathbf{r}_p - \mathbf{r}|}. \qquad \ldots\ldots(2.3)$$

$\mathbf{j}_n(\mathbf{r})$ is here the nuclear current density, which will contain contributions from both convection currents and intrinsic spins. $V_A^{(1)}$ is sufficient for electrons, if we use the Dirac theory with

$$\mathbf{v}_p = c\boldsymbol{\alpha}, \qquad \ldots\ldots(2.4)$$

but for a nuclear projectile (treated non-relativistically) we add the interaction $V_A^{(2)}$ of the projectile's magnetic moment $\boldsymbol{\mu}_p$ (in nuclear magnetons) with the magnetic field of the nucleus:

$$V_A^{(2)} = -\frac{e\hbar}{2Mc}\boldsymbol{\mu}_p \cdot \mathrm{curl}_p\, \mathbf{A}(\mathbf{r}_p), \qquad \ldots\ldots(2.5)$$

where M is the proton mass. $V_A^{(1)}$ and $V_A^{(2)}$ lead to magnetic transitions, as well as giving a correction to the electric transitions induced by V. These formulae suggest at once why the magnetic interaction is generally much weaker than the electric. $V_A^{(1)}$, in comparison with V, contains a factor v_p/c and also (implicitly from $\mathbf{j}_n$) a factor v_k/c, where v_k represents the velocity of a nucleon in the nucleus. The first factor will be small in the case of nuclear projectiles, and while this factor will be of order unity for relativistic electrons, the second factor will always operate to keep down magnetic effects. Consideration of $V_A^{(2)}$ (in the case of nuclear projectiles) does not alter this conclusion: a comparison of dimensional factors in (2.2) and (2.5) shows

$$\frac{V_A^{(2)}}{V_A^{(1)}} \sim \frac{\hbar\mu_p}{Mr_p v_p} \sim \frac{\text{spin angular momentum of projectile}}{\text{orbital angular momentum of projectile}}$$

and this will generally be small, as the projectile passing outside the nucleus has usually large orbital angular momentum.

So far we have regarded the electromagnetic field as classical and static. More properly it should be treated by quantum radiation theory (Jean and Prentki 1954, Biedenharn, McHale and Thaler 1955, Alder, Bohr *et al.* 1956, § II.B, Schiff 1954). The interaction of the total field $\mathbf{A}, \phi$ with the particles is then given by

$$\int \left(\rho\phi - \frac{\mathbf{j}\cdot\mathbf{A}}{c}\right) d\tau$$

where ρ and $\mathbf{j}$ are the total charge and current density, a sum of contributions from the projectile and the nucleus. This is used to calculate the transition probability from an initial state ψ_i of the nucleus, ϕ_i of the projectile to a final state ψ_f, ϕ_f in which energy ΔE has been lost by the projectile and gained by the nucleus. This is a second-order radiative process, involving intermediate states in which a virtual photon exists in transit from the one body to the other. It turns out that the result of the calculation is exactly the same as if we eliminate the radiation field from explicit consideration, and calculate the transition between the states of the projectile–nucleus system by means of an effective interaction Hamiltonian only

slightly modified from the above interaction $(V+V_A^{(1)}+V_A^{(2)})$†; the modification consists in replacing the function $|\mathbf{r}_p-\mathbf{r}_k|^{-1}$ in (2.1) and (2.3) by the retarded Green's function

$$\frac{1}{|\mathbf{r}_p-\mathbf{r}_k|} \to \frac{\exp(i\kappa|\mathbf{r}_p-\mathbf{r}_\kappa|)}{|\mathbf{r}_p-\mathbf{r}_\kappa|}, \qquad \ldots\ldots(2.6)$$

where

$$\kappa = \Delta E/\hbar c, \qquad \ldots\ldots(2.7)$$

i.e. κ is the wave number of a virtual photon which conserves energy ($\sim$ a real photon). This can be interpreted by supposing that, while the interactions (2.1), (2.2) and (2.5) with (2.3) would be correct for a static nuclear distribution of charge and current, the nuclear distribution involved in the transition is oscillating with frequency $\Delta E/\hbar$, and so the potentials due to it are given by just the classical formulae for retarded potentials of this frequency. The retardation correction vanishes with vanishing excitation energy.

It is more customary, in calculating the excitation of the nucleus, to think of the interaction inversely, as the action of the potentials (scalar and vector) produced by the projectile on the charge and current distribution of the nucleus; however, this would evidently lead to the same interactions (2.1), (2.2) and (2.5). Only, by considering instead the action of the potential due to the nucleus on the projectile, it becomes more understandable that the retardation correction is of the simple form (2.6), however high the velocity of the projectile. While the electrostatic potential (2.1) may in principle be put in the Schrödinger equation for the particles and an exact wave function then sought, the retarded potentials are only valid for calculating transitions in first order perturbation, although as a better approximation the electrostatic part V (2.1) may be split off from the total retarded interaction and treated exactly, and the remainder treated in first order perturbation only (equivalent to use of ' Coulomb gauge ', Heitler 1954).

Although the retardation effect has been dealt with in calculations by the authors mentioned above, in all practical applications so far it has been neglected, i.e. interactions (2.1) and (2.2), (2.5) with (2.3) have been used. (2.6) shows the condition for the retardation effect to be negligible:

$$r_{p,\max} \ll \lambda\!\!\!^{-}_{\text{photon}}, \qquad \ldots\ldots(2.8)$$

where $r_{p,\max}$ is the maximum distance of the projectile from the nucleon at which the excitation occurs, and $\lambda\!\!\!^{-}_{\text{photon}} = 1/\kappa$. This condition will be investigated further in § 5.3 and § 6.3.

The electrostatic interaction (2.1) can be expanded as a series of multipole terms:

for $r_k < r_p$:

$$V = 4\pi Z_1 e^2 \sum_{k,\lambda,\mu} \frac{1}{2\lambda+1} \frac{r_k^{\lambda}}{r_p^{\lambda+1}} Y_{\lambda\mu}(\theta_p,\phi_p) Y_{\lambda\mu}{}^*(\theta_k,\phi_k); \qquad \ldots\ldots(2.9)$$

for $r_k > r_p$:

$$V = 4\pi Z_1 e^2 \sum_{k,\lambda,\mu} \frac{1}{2\lambda+1} \frac{r_p^{\lambda}}{r_k^{\lambda+1}} Y_{\lambda\mu}(\theta_p,\phi_p) Y_{\lambda\mu}{}^*(\theta_k,\phi_k). \qquad \ldots\ldots(2.10)$$

† The last term $V_A^{(2)}$ is to be omitted for Dirac electrons.

For nuclear projectiles it is generally assumed that Coulomb excitation occurs only when the projectile is outside the nucleus, so only form (2.9) need be considered. This is not true for electrons. A λ term in V can lead to excitation of an electric transition Eλ in the nucleus (i.e. vector change of nuclear angular momentum by $\hbar\lambda$, change of parity $(-1)^\lambda$). λ runs from 0 to ∞, but in (2.9) the $\lambda = 0$ term is a constant in the coordinates of particle k and so causes no transitions, whereas the $\lambda = 0$ term in (2.10) is proportional to r_k^{-1} and can cause transitions. This means that electric monopole transitions (0–0, no parity change) can be produced only by particles, such as electrons, which penetrate the nucleus.

A similar but more complicated multipole expansion can be performed on the interactions $V_A^{(1)}$ and $V_A^{(2)}$ giving multipole operators of order $\lambda(1-\infty)$, which can induce magnetic transitions Mλ in the nucleus (i.e. vector change of nuclear angular momentum $\hbar\lambda$, change of parity $(-1)^{\lambda+1}$), and giving additional terms in the electric transitions.

§ 3. Approximate Methods

We discuss here the principles of the methods of calculation used, and develop them in more detail later.

3.1. *Semi-classical Method*

The projectile is treated as a classical particle, moving on an orbit past the nucleus. The latter, on the other hand, is treated according to wave mechanics. The Coulomb potential (2.1) is then a time-dependent external field $V(t)$ acting on the nucleus, determined by the instantaneous position $\mathbf{r}_p(t)$ of the projectile. Transitions from an initial state of the nucleus ψ_i to a final state ψ_f are evaluated by time-dependent perturbation theory. The well-known result for the eventual amplitude b_{fi} of the state ψ_f is, in first order,

$$b_{fi} = -\frac{i}{\hbar}\int_{-\infty}^{\infty} \langle\psi_f|\, V(t)\,|\psi_i\rangle\, e^{i\omega t}\, dt, \qquad \ldots\ldots(3.1)$$

where

$$\omega = \Delta E/\hbar = (E_f - E_i)/\hbar. \qquad \ldots\ldots(3.2)$$

The probability of excitation by the projectile on this orbit is $|b_{fi}|^2$, and so the element of cross section contributed by particles in orbits with impact parameters p–$(p+dp)$ (figure 1) and azimuths ϕ–$(\phi+d\phi)$ is

$$d\sigma = |b_{fi}|^2\, p\, dp\, d\phi. \qquad \ldots\ldots(3.3)$$

This is essentially the method first used for calculating atomic excitation by fast particles by Gaunt (1927) (see also Bohr 1948). The total cross section is obtained by integrating over impact parameters p. It is necessary to investigate the orbits more carefully.

One treatment (' method of Rutherford orbits ') is to separate out from the Coulomb potential (2.1) the point charge interaction between Z_1 and Z_2

$$V_0 = Z_1 Z_2 e^2/r_p, \qquad \ldots\ldots(3.4)$$

and to employ the Rutherford-scattering classical orbits of the projectile under this force, neglecting any disturbance of the orbit due to excitation of the nucleus.

Then the potential remaining to cause excitation in (3.1) is not V but

$$V' = V - V_0. \qquad \ldots\ldots(3.5)$$

This procedure is only practicable for non-relativistic projectiles (nuclear projectiles). The calculation was first carried through by Ter-Martirosyan (1952). The method is quite exact (to first order in V') provided

(i) $$\Delta E \ll T, \qquad \ldots\ldots(3.6)$$

where T is the kinetic energy of the projectile. This ensures that the loss of energy of the projectile does not seriously alter its motion ;

(ii) $$\eta \gg 1, \qquad \ldots\ldots(3.7)$$

where $$\eta = Z_1 Z_2 e^2/\hbar v, \qquad \ldots\ldots(3.8)$$

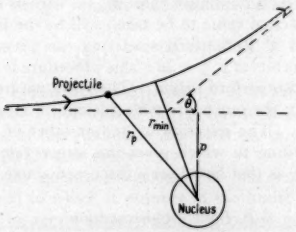


Figure 1. Rutherford orbit of projectile.

v being the value of v_p far from the nucleus. This is the condition for the validity of classical mechanics in Rutherford scattering, and has been much discussed (e.g. Bohr 1948, § 1.3, Mott and Massey 1949, § 7.5). It may be understood by comparing the classical closest distance of approach $2a$, defined by

$$a = Z_1 Z_2 e^2/m_p v^2 \qquad \ldots\ldots(3.9)$$

with the wavelength λ of the projectile :

$$\lambda = 1/k = \hbar/m_p v.$$

We see that $$\eta = a/\lambda = ak, \qquad \ldots\ldots(3.10)$$

and so condition (3.7) is equivalent to $a \gg \lambda$. Thus classical mechanics is valid provided the wave nature of the projectile is inconspicuous at the position of closest approach. Condition (3.7) is the better satisfied by reducing the velocity of the projectile, but this is limited by requirement (3.6). In the opposite limit $\eta \ll 1$ a quantum treatment of the collision by Born approximation is valid.

The differential cross section as well as the total may be obtained by the Rutherford orbital method. It is possible to assume that a projectile which causes excitation continues on its Rutherford orbit and is thus scattered through the Rutherford angle. The differential cross section is therefore

$$\sigma(\theta, \phi) = \sigma_R(\theta) |b_{if}|^2. \qquad \ldots\ldots(3.11)$$

where $\sigma_R(\theta)$ is the Rutherford cross section

$$\sigma_R(\theta) = a^2/4\sin^4\tfrac{1}{2}\theta, \qquad \ldots\ldots(3.12)$$

and b_{if} refers to the appropriate orbit.

The recoil of the nucleus may be taken into account exactly : we simply replace the mass m_p by the reduced mass of projectile and nucleus m_0 ; also $\mathbf{r}_p$ and the nuclear internal coordinates $\mathbf{r}_k$ are to be measured from the mass centre of the nucleus.

The semi-classical approach still has some use for explanatory purposes even when the full Rutherford orbital method is not valid owing to failure of (3.7). Sufficiently distant orbits may always be treated classically, and we may approximate the orbits to straight lines (Huby and Newns 1951). Then the cross section is calculated according to (3.1) and (3.3), but only with integration over impact parameters beyond a certain minimum value p_{min}. Various effects lead to different values of p_{min}, and the actual value to be taken will be the largest of these. Thus the effect on the orbits of Rutherford scattering can be crudely represented by cutting off the straight orbits at $p_{min} = 2a$: this procedure is approximately equivalent to the method of Rutherford orbits. The wave nature of the projectile calls for a cut-off at $p_{min} = \lambda$: the procedure with this cut-off is very roughly equivalent to Born approximation. The selection of one or other of these methods, under given conditions, according to which gives the larger value of p_{min}, is evidently approximately the same as that based on η discussed above.

3.2. *Born Approximation*

If the projectile and nucleus are both treated wave-mechanically, the first-order excitation due to the interaction V (2.1) is given by Born approximation. The transition takes the nucleus from ψ_i to ψ_f and the projectile from a plane wave of momentum $\hbar\mathbf{k}_i$ to another $\hbar\mathbf{k}_f$. The differential cross section is

$$\sigma(\theta,\phi) = \frac{k_f^2}{4\pi^2\hbar^2 v_i v_f}|\langle \mathbf{k}_f, \psi_f | V | \mathbf{k}_i, \psi_i\rangle|^2. \qquad \ldots\ldots(3.13)$$

The plane waves are normalized as $\exp(i\mathbf{k}.\mathbf{r}_p)$. In a non-relativistic case, nuclear recoil is exactly taken into account in the usual way by measuring all the coordinates $\mathbf{r}_p$ and $\mathbf{r}_k$ from the centre of mass of the target nucleus, and using the reduced mass m_0. A non-relativistic nuclear projectile can be treated simply as a point charge without spin, even though it may actually have a spin, since the latter is unaffected by the Coulomb potential V.

For a relativistic electron the Dirac theory must be used for the free-electron wave functions.

As mentioned in § 3.1, the condition for validity of Born approximation in Coulomb scattering is $\eta \ll 1$ (for both η_i and η_f). For relativistic electrons we have from (3.8)

$$\eta \sim Z_2/137, \qquad \ldots\ldots(3.14)$$

and so Born approximation is good for light target nuclei. On the other hand, for nuclear projectiles of energy well below the Coulomb barrier, η is generally large (§ 7.2), and Born approximation is useless.

3.3. *Method of Coulomb-distorted Waves*

The reason for the failure of Born approximation, i.e. the treatment of the projectile waves as only slightly perturbed in collision, when η is large, is principally that the elastic Rutherford scattering is then strong: indeed this can be described by means of classical orbits as in § 3.1. The remedy is to separate out from V the point charge interaction V_0 between Z_1 and Z_0 (3.4), and treat this exactly; leaving the remainder V' (3.5) to be dealt with by perturbation theory. The projectile states before and after collision are now exact Coulomb wave functions $\phi(\mathbf{k}, \mathbf{r}_p)$ for a particle in the potential V_0, and having momentum $\hbar\mathbf{k}$ at infinity (see for example Mott and Massey 1949, Chapter III). The cross section for transitions due to V' resulting in $\mathbf{k}_i \to \mathbf{k}_f$ for the projectile and $\psi_i \to \psi_f$ for the nucleus is obtained by the method of distorted waves (Mott and Massey 1949, § 8.5). The formula (to lowest order in V) is the same as for the Born approximation (3.13), except for the substitution of new wave functions and potential:

$$\sigma(\theta, \phi) = \frac{k_f^2}{4\pi^2 \hbar^4 v_i v_f} |\langle \bar{\phi}_f, \psi_f | V' | \phi_i, \psi_i \rangle|^2. \qquad \ldots\ldots(3.15)$$

Here the ϕ's are again normalized as $\exp(i\mathbf{k}\mathbf{r}_p)$. $\bar{\phi}_f$ in the matrix element means that the latter is calculated not with $\phi^*(\mathbf{k}_f, \mathbf{r}_p)$ but with $\phi(-\mathbf{k}_f, \mathbf{r}_p)$ instead; this is a well-known feature of the method of distorted waves. The remarks on Born approximation below (3.13) still apply.

This method is the quantum analogue of the semi-classical method of Rutherford orbits (§ 3.1) to which it reduces in the limits of conditions (3.6) and (3.7) and it may thus be regarded as yielding quantum corrections to the earlier method. It has been extensively worked out for nuclear projectiles.

§ 4. Order-of-Magnitude Estimates

4.1. *Formulae for Total Cross Section*

We use the method of straight orbits outlined in § 3.1. The transition amplitude b_{if} is given by (3.1) together with the potential V in multipole expansion (2.9), which is adequate as regards order of magnitude for all electric excitations except monopole excitation by electrons, and this is deferred to § 6.1. For given initial and final nuclear states of spins I_i, I_f, the matrix element $\langle \psi_f | V | \psi_i \rangle$ selects from V only the multipoles λ which can effect the required change of angular momentum and parity, and normally only one value of λ (the lowest permitted) need be considered. The result is

$$b_{if} = -\frac{4\pi i Z_1 e^2}{\hbar(2\lambda+1)} S_{E\lambda,(M_i - M_f)} \times \left\langle I_f M_f \middle| \sum_k r_k^{\lambda} Y_{\lambda(M_i - M_f)}^*(\theta_k, \phi_k) \middle| I_i M_i \right\rangle, \qquad \ldots\ldots(4.1)$$

where

$$S_{E\lambda,\mu} = \int_{-\infty}^{\infty} \frac{e^{i\omega t} Y_{\lambda\mu}[\theta_p(t), \phi_p(t)]}{r_p^{\lambda+1}} dt. \qquad \ldots\ldots(4.2)$$

The orbital integral $S_{E\lambda,\mu}$ is essentially the Fourier component with transition frequency ω of the multipole potential pulse acting on the nucleus. To estimate it, we note that the spherical harmonic is not a vital factor, and may be roughly

replaced by an average value $(2\lambda+1)^{-\frac{1}{2}}$. The integrand has a peak value of $(2\lambda+1)^{-\frac{1}{2}}p^{-(\lambda+1)}$, and the pulse a duration of approximately p/v, the time of passage of the projectile past the nucleus. We write

$$\xi(p) = \omega p/v. \qquad \ldots\ldots(4.3)$$

Then if $\xi(p) < 1$ the exponential factor in (4.2) is approximately constant throughout the pulse, and we may take $S_{\mathrm{E}\lambda,\,\mu} \sim$ peak value $\times$ duration of pulse

$$\sim (2\lambda+1)^{-\frac{1}{2}}p^{-\lambda}v^{-1}. \qquad \ldots\ldots(4.4)$$

If $\xi(p) > 1$, the exponential factor in (4.2) oscillates rapidly during the pulse, and $S_{\mathrm{E}\lambda,\,\mu}$ is much reduced: the nuclear states adjust themselves adiabatically to the slow pulse, giving little final excitation. In fact, for large $\xi(p)$, $S_{\mathrm{E}\lambda,\,\mu}$ can be shown to fall off as $\exp\{-\xi(p)\}$ (Alder, Bohr *et al.* 1956, § II.E.7). Substituting (4.4) in (4.1) we obtain for the value of $|b_{\mathrm{If}}|^2$ averaged over M_{I} and summed over M_{f}

$$\overline{|b_{\mathrm{If}}|^2} = \frac{1}{(2\lambda+1)^3}\left(\frac{4\pi Z_1 e^2}{\hbar v\, p^{\lambda}}\right)^2 B(\mathrm{E}\lambda), \qquad \ldots\ldots(4.5)$$

where $B(\mathrm{E}\lambda)$ is the reduced radiative transition probability†

$$B(\mathrm{E}\lambda) = \frac{1}{(2I_{\mathrm{I}}+1)}\sum_{M_{\mathrm{I}},\,M_{\mathrm{f}},\,\mu}\left|\left\langle I_{\mathrm{f}}\,M_{\mathrm{f}}\left|\sum_k r_k^{\lambda}\,\mathrm{Y}_{\lambda\mu}(\theta_k,\phi_k)\right|I_{\mathrm{I}}\,M_{\mathrm{I}}\right\rangle\right|^2. \qquad \ldots\ldots(4.6)$$

To obtain the total cross section for this transition, (4.5) is substituted in (3.3) and integrated over impact parameters. The result is

for $\lambda = 1$:
$$\sigma_{\mathrm{E1}} = \frac{32\pi^3}{27}\left(\frac{Z_1 e^2}{\hbar v}\right)^2 B(\mathrm{E1})\ln\left(\frac{p_{\max}}{p_{\min}}\right); \qquad \ldots\ldots(4.7)$$

for $\lambda > 1$:
$$\sigma_{\mathrm{E}\lambda} = \frac{16\pi^3}{(2\lambda+1)^3(\lambda-1)}\left(\frac{Z_1 e^2}{\hbar v}\right)^2 \frac{B(\mathrm{E}\lambda)}{p_{\min}^{2\lambda-2}}. \qquad \ldots\ldots(4.8)$$

As explained in § 3.1, it has been necessary to cut off the integration at a lower impact parameter $p_{\min}$. There is also an upper cut-off at $p_{\max}$, the distance at which the excitation becomes adiabatic, i.e. $\xi(p_{\max}) = 1$, as given by (4.3); hence

$$p_{\max} = v/\omega. \qquad \ldots\ldots(4.9)$$

However, in deriving (4.8) the integration has been continued to infinity, since the integral converges quickly: this is valid provided $p_{\max} \gg p_{\min}$, i.e. not all orbits are adiabatic. In the dipole case (4.7), on the other hand, contributions from distant orbits are more important, and the integral does not converge unless the cut-off is inserted explicitly at $p_{\max}$.

In the case of interest for nuclear projectiles, $\eta \gg 1$, we put $p_{\min} = 2a$, and so finally:

for $\lambda = 1$:
$$\sigma_{\mathrm{E1}} = \frac{32\pi^3}{27}\left(\frac{Z_1 e^2}{\hbar v}\right)^2 B(\mathrm{E1})\ln\left(\frac{v}{2a\omega}\right); \qquad \ldots\ldots(4.10)$$

for $\lambda > 1$:
$$\sigma_{\mathrm{E}\lambda} = \frac{16\pi^3}{(2\lambda+1)^3(\lambda-1)\,2^{2\lambda-2}}\left(\frac{Z_1 e^2}{\hbar v}\right)^2 \frac{B(\mathrm{E}\lambda)}{a^{2\lambda-2}}. \qquad \ldots\ldots(4.11)$$

† This definition differs from that of Alder, Bohr *et al.* (1956, p. 436) inasmuch as we have removed the electric charge from the matrix element.

The requirement that not all orbits be adiabatic can be expressed in terms of a parameter ξ (cf. 4.3).

$$\xi = \frac{\omega a}{v} = \eta \frac{\Delta E}{2T} = \frac{Z_1 Z_2 m_0^{1/2} e^2 \Delta E}{\hbar (2T)^{3/2}}. \qquad \ldots\ldots (4.12)$$

The requirement then is roughly $\xi < 1$, which together with $\eta > 1$ presents the conditions of approximate validity of (4.10) and (4.11) (these conditions evidently imply also (3.6)). Formulae (4.10) and (4.11) are a good approximation to the accurate result by the Rutherford orbital method § 5.1, under the stated conditions. If ξ becomes large, it follows from above remarks that the cross section will be reduced from (4.10) and (4.11) by an exponential factor, which accurate calculation (Ter-Martirosyan 1952) shows to be of the order of $\exp(-2\pi\xi)$.

In the régime of Born approximation, $\eta \ll 1$, we can try using the cross section formulae (4.7) and (4.8), but putting $p_{\min} \sim \lambda \sim 1/k$. This indeed gives formulae agreeing in form with those obtained by Born approximation, but the numerical factors are incorrect.† We therefore anticipate the Born approximation results for relativistic electron scattering of § 6.1. Using (6.1)–(6.7) we find for the total cross section, when $T \gg \Delta E$,

for $\lambda = 1$:
$$\sigma_{E1} = \frac{32\pi^2}{9} \left(\frac{e^2}{\hbar c}\right)^2 B(E1) \ln\left(\frac{2kc}{\omega}\right); \qquad \ldots\ldots (4.13)$$

for $\lambda > 1$:
$$\sigma_{E\lambda} = \frac{2^{2\lambda+2} \pi^2 k^{2\lambda-2}}{[(2\lambda+1)!!]^2 \lambda(\lambda-1)} \left(\frac{e^2}{\hbar c}\right)^2 B(E\lambda). \qquad \ldots\ldots (4.14)$$

The straight-orbit requirement that not all orbits be adiabatic is $\xi(1/k) \ll 1$, which turns out to be equivalent to $\Delta E \ll T$. There is a further requirement for the validity of (4.13) and (4.14) in addition to $\eta \ll 1$, $\Delta E \ll T$, viz. that kR_0 (R_0 radius of nucleus) be not much larger than unity: for in that case, the excitation takes place principally inside the nucleus, so (2.9) is a bad approximation to the interaction, and the cross section is reduced owing to diffraction of the projectile waves within the nucleus. When the requirement $\Delta E \ll T$ is violated, the cross sections fall below (4.13) and (4.14).

4.2. *Discussion of Total Cross Section*

The reduced radiative transition probabilities $B(E\lambda)$ must be considered. We refer these to the ' single-particle ' estimates of Blatt and Weisskopf (1952, § 12.6).

$$B_{sp}(E\lambda) = \frac{(2\lambda+1)}{4\pi} \left(\frac{3}{3+\lambda}\right)^2 R_0^{2\lambda}, \qquad \ldots\ldots (4.15)$$

where R_0 is the nuclear radius, estimated at $1{\cdot}5A^{1/3} \times 10^{-13}$ cm; but we bear in mind that for many of the low-lying E2 transitions which have proved most interesting $B(E2)$ is from 10 to 100 times the single particle value, owing to collective effects, whereas low-energy dipole transitions are not only very scarce but such as do exist have values of $B(E1)$ much smaller than the single particle value (e.g. Goldhaber and Weneser 1955). To compare the role of $B(E\lambda)$ here with its

† We actually get good numerical agreement by inserting $p_{\min} = 2/k$ in (4.7) and (4.8), but this is arbitrary.

role in radiative transitions, we note that the probability per unit time for the inverse decay of state f of the nucleus to state i, with emission of a photon of multipole order λ (electric or magnetic), is

$$T_\lambda = \frac{8\pi(\lambda+1)}{\lambda[(2\lambda+1)!!]^2}\cdot\frac{e^2}{\hbar}\kappa^{2\lambda+1}\frac{(2I_i+1)}{(2I_f+1)}B(\lambda). \qquad \ldots\ldots(4.16)$$

We can make an instructive comparison of the excitation cross section with the geometrical cross section of the nucleus πR_0^2, by substituting (4.15) in (4.10)–(4.14). We first consider nuclear projectiles. The ratio of the cross section to the geometrical is the product of three factors:

(i) $(Z_1 e^2/\hbar v)^2$. This is usually rather small, reflecting the electromagnetic nature of the problem. Thus for protons of 2–10 MeV it varies from 1/80 to 1/400; and for α's of the same energy from 1/5 to 1/25.

(ii) $(R_0/a)^{2\lambda-2} = (2T/E_B)^{2\lambda-2}$, where E_B is the Coulomb barrier:

$$E_B = Z_1 Z_2 e^2/R_0 \sim Z_1 Z_2 A_2^{-1/3}\ \text{MeV}. \qquad \ldots\ldots(4.17)$$

This factor will be unity when the energy T is half the barrier, and it increases with energy (except for $\lambda = 1$). An energy of half the barrier height is within the practicable region for many nuclei and an energy approaching the barrier is as high as is at all desirable. In the dipole case there is a further logarithmic factor, which is of lesser importance.

(iii) Numerical factors N_λ, which amount to

λ	1	2	3	4
N_λ	1·6	0·045	$2{\cdot}0\times10^{-3}$	$1{\cdot}5\times10^{-4}$

Thus altogether, for protons of energy half the Coulomb barrier, the single particle estimate of the cross section is 10^{-4}–10^{-3} of the geometrical cross section; and the cross section falls off by roughly a power of ten for each unit increase of λ. On account of factor (ii), the higher multipoles are favoured by increased energy of the projectile. The excitation function according to these formulae is proportional to $T^{2\lambda-3}$, i.e. rising with energy for all but the dipole case, but even in this case it turns out that, owing to the logarithmic factor, the cross section rises slowly with energy until ξ becomes extremely small. With decreasing energy, the cross section will fall off much more steeply than our formulae show, when ξ becomes greater than unity (see figure 7).

The case of the electron as projectile can be considered similarly. The ratio of the cross section given by (4.13), (4.14) to the geometrical is the product of three factors:

(i) $(e^2/\hbar c)^2 = (1/137)^2$. This 'electromagnetic factor' is much smaller than the factor (i) for nuclear projectiles, because of the higher velocity.

(ii) $$(kR_0)^{2\lambda-2} = (TR_0/\hbar c)^{2\lambda-2} \sim (T_{\text{MeV}}\, A_2^{1/3}/130)^{2\lambda-2}. \qquad \ldots\ldots(4.18)$$

This depends on the ratio of the nuclear radius to the electron's wavelength $\lambda\!\!\!^{-} = 1/k$. It is unity when $R_0 = \lambda\!\!\!^{-}$, and it increases with energy (except for $\lambda = 1$). The condition $R_0 = \lambda\!\!\!^{-}$ is well within the practicable energy region (e.g. with $A_2 = 12$, it requires $T \sim 60$ MeV), and energies several times greater (~ 200 MeV) have been much used.

In the dipole case there is a further logarithmic factor of lesser importance.

(iii) Numerical factors N_λ' :

λ	1	2	3	4
N_λ'	1·5	0·064	$1{\cdot}7\times10^{-3}$	$4{\cdot}0\times10^{-5}$

Thus altogether, for electrons with $R_0\sim\lambda$, the single-particle estimate of the E2 cross section is about 3×10^{-6} of the geometrical cross section, and the cross section falls off by about a factor of 30 for each unit increase of λ. Higher multipoles are favoured by increased energy of the projectile. The excitation function is proportional to $T^{2\lambda-2}$, i.e. rising with energy. However, if kR_0 rises much above unity, then the increase of cross section with energy slows down and finally turns to a decrease, owing to diffraction.

The region of energy for which $R_0\sim\lambda$ is in fact the most desirable one for the purpose of ' probing ' the nuclear structure, because there the detailed angular distribution and total cross section depend on the shape of the nuclear wave functions (§ 6.1).

We see that there is a close parallelism between the cases of nuclear and electronic projectiles. It is of interest that an increase of electric multipole order diminishes the Coulomb excitation cross section to a much lesser extent (under favourable conditions) than it decreases the rate of radiative decay in the inverse transition. This can be seen from (4.16) which contains a factor $(\kappa R_0)^{2\lambda}$ in place of the factor (ii) for Coulomb excitation, viz. $(R_0/a)^{2\lambda-2}$ or $(kR_0)^{2\lambda-2}$. We shall normally have $\kappa\ll a^{-1}$ and $\kappa\ll k$. Therefore on increasing λ the extra factor introduced is much smaller in the case of photon emission than of excitation by particles. Similarly, magnetic transitions can compete with electric in radiative decay but not in Coulomb excitation, because the small factor v_k/c in magnetic moments may be off-set in decay when the magnetic multipole order is lower than the competing electric (e.g. M1 against E2), owing to the role of the κR_0 factors, whereas there are no strong factors like the latter in Coulomb excitation to militate against the larger λ ; but on the contrary there is an additional factor v_p/c to reduce magnetic excitation relative to electric (for non-relativistic projectiles).

4.3. *Angular Distribution of Inelastically Scattered Projectiles*

It is possible to explain qualitatively also the angular distribution of the inelastically scattered projectiles by a refinement of the method of straight orbits : in its crude form this would always predict purely forward scattering.

We consider first the case of nuclear projectiles under the conditions of approximate validity of the total cross section formulae (4.10), (4.11), $\xi<1$, $\eta>1$. Angular distributions for $\eta\gg1$ calculated by the exact method of Rutherford orbits, for various values of ξ, are shown in figure 2. Let us take first the case of $\xi=0$ (all orbits non-adiabatic). The cross section is accurately given by (3.11) and (3.12) but we approximate this by using for $|b_{if}|^2$ the straight-orbit formula (4.5) calculated with the value of the impact parameter p which actually leads to Rutherford scattering through an angle θ, viz.

$$p=a\cot\tfrac{1}{2}\theta. \qquad \ldots\ldots(4.19)$$

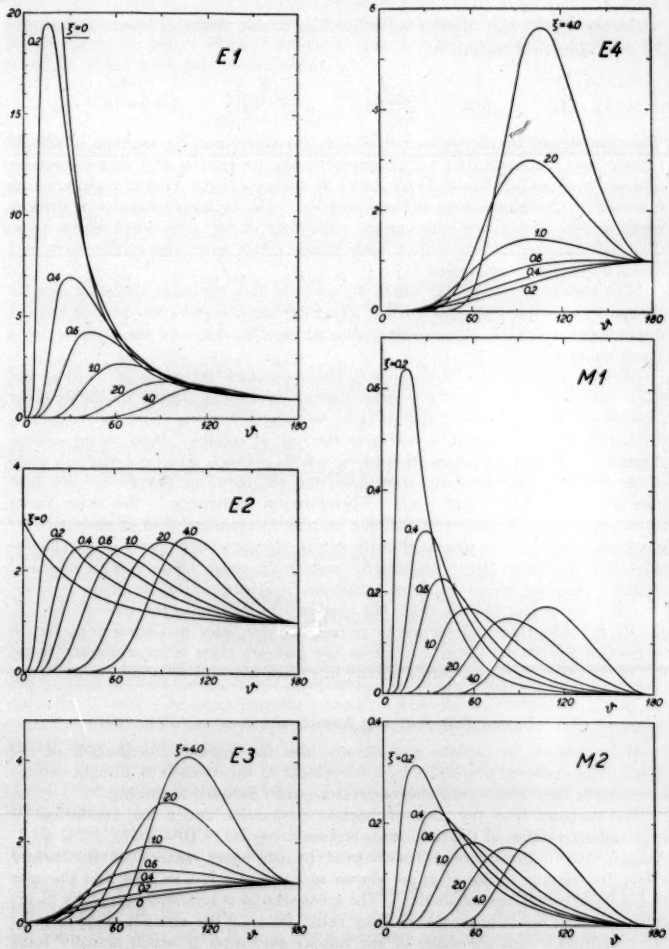


Figure 2. Angular distribution of inelastically scattered projectiles in semi-classical approximation. The ordinate is $df_\lambda(\theta, \xi)/d\Omega$, and the abscissa is θ. Each curve corresponds to a fixed value of ξ. The Eλ cross sections are normalized to unity at 180, while the magnetic are normalized to give a total cross section of unity. (*After* Alder, Bohr *et al.* 1956.)

A better approximation than (4.5) for small impact parameters is to replace p in this formula by the actual distance of closest approach on the Rutherford orbit defined by θ:

$$r_{\min} = a(1+\sin\tfrac{1}{2}\theta)/\sin\tfrac{1}{2}\theta.$$

Then from (3.11), (3.12), and (4.5) we obtain

$$\sigma_{\mathrm{E}\lambda}(\theta) \propto \frac{(\sin\frac{1}{2}\theta)^{2\lambda-4}}{(1+\sin\frac{1}{2}\theta)^{2\lambda}}. \qquad \ldots\ldots(4.20)$$

This simple formula represents fairly well the actual trend of the Eλ, $\xi = 0$ curves in figure 2. The general idea may be summarized thus: the larger λ, the more important are close collisions (because of the factor $r_{\mathrm{p}}^{-(\lambda+1)}$ in V (2.9)), and hence backward Rutherford scattering becomes more important.

In the case of finite ξ we shall expect the same result for the closer, non-adiabatic orbits (corresponding to large θ); but the more distant orbits (corresponding to small θ) will be adiabatic, with $\xi(p) > 1$ (4.3), and so the cross section will be reduced by an exponential factor. This explains the drop in the angular distributions of figure 2 at small angles. We find (from (4.3), (4.12) and (4.19)) that the scattering angle $\theta_{\min}$ below which $\xi(p)$ becomes large, and so the cross section drops rapidly, is proportional to ξ. This is approximately borne out by figure 2, which suggests $\theta_{\min} \sim \xi$. The hump in the angular distributions at intermediate angles, growing with increasing ξ, is not readily explainable.

It is useful to have an estimate of the ratio of the differential cross section to that for Rutherford scattering, i.e. $|b_{\mathrm{if}}|^2$. From the improved formula (4.5), using the single-particle estimate of $B(\mathrm{E}\lambda)$ (4.15) we find the formula, approximately valid for small ξ and large η, θ:

$$\frac{\sigma_{\mathrm{E}\lambda}(\theta)}{\sigma_{\mathrm{R}}(\theta)} = N_\lambda''\left(\frac{Z_1 e^2}{\hbar v}\right)^2\left(\frac{R_0}{a}\right)^{2\lambda}\left(\frac{\sin\frac{1}{2}\theta}{1+\sin\frac{1}{2}\theta}\right)^{2\lambda}, \qquad \ldots\ldots(4.21)$$

where N_λ'' has the values

λ	1	2	3	4
N_λ''	0·79	0·18	0·06	0·03

The factors here, apart from the last, are similar in form and behaviour to those expressing the ratio of the total cross section to geometrical, discussed in § 4.2.

We turn to the régime of Born approximation, i.e. high energy electrons, with $\eta < 1$, $T \gg \Delta E$, for which the total cross section is given by (4.13), (4.14). Scattering of a particle through an angle θ means a change of momentum by

$$\begin{aligned}\hbar K &= \hbar|\mathbf{k}_{\mathrm{i}} - \mathbf{k}_{\mathrm{f}}| \\ &= \hbar(k_{\mathrm{f}}^2 + k_{\mathrm{i}}^2 - 2k_{\mathrm{i}}k_{\mathrm{f}}\cos\theta)^{\frac{1}{2}} \sim 2\hbar k\sin\tfrac{1}{2}\theta. \qquad \ldots\ldots(4.22)\end{aligned}$$

By the uncertainty principle, this change of momentum must be associated with scattering over a spatial region of dimensions of the order of $1/K$. Thus we may crudely correlate scattering through θ in Born approximation with excitation as given by the theory of straight orbits, with an impact parameter

$$p \sim 1/K. \qquad \ldots\ldots(4.23)$$

The differential cross section can now be obtained from (3.3) with (4.5), (4.22) and (4.23):

$$\sigma(\theta) \propto K^{2\lambda-4} \propto (\sin \tfrac{1}{2}\theta)^{2\lambda-4}. \qquad \ldots\ldots(4.24)$$

This is rather similar to the previous result (4.20) and has the same general interpretation: for larger λ, close collisions are more important and these are associated with large angle scattering (though now this is due to the uncertainty principle), and so backward scattering becomes more important.

To discuss Born approximation by semi-classical methods as we have done is evidently not very reliable. The actual relativistic Born approximation result § 6.1 differs in various ways from our estimate (4.24), but it contains the same dominating factor. The ratio of the differential cross section to that for scattering by a point charge Z_2, generally called $|F|^2$ in the literature on electrons, could be obtained by the procedure used to calculate (4.24), but this yields incorrect numerical factors, and so we shall anticipate the exact result (6.5). This, together with the single-particle estimate (4.15), yields, on omitting a diffraction factor which is only important for large KR_0 and which causes a decrease at larger angles,

$$|F_{E\lambda}|^2 = \frac{N_\lambda''}{Z_2^2} (kR_0)^{2\lambda} (\sin \tfrac{1}{2}\theta)^{2\lambda}, \qquad \ldots\ldots(4.25)$$

where N_λ'' has the values

λ	1	2	3	4
N_λ''	0·75	0·13	0·01	5×10^{-4}

The factors here, apart from the last, are similar in form and behaviour to those expressing the ratio of the total cross section to the geometrical, discussed in § 4.2, except that now the ratio contains a factor Z_2^{-2} in place of $(e^2/\hbar c)^2$.

§ 5. Excitation by Nuclear Projectiles

We review here the exact calculations which have been made for the case of nuclear projectiles.

5.1. *Semi-classical Method* (*method of 'Rutherford orbits'*)

The method, applicable when $\eta \gg 1$, $\Delta E \ll T$, has been explained in § 3.1. We consider first the total cross section which is obtained by integrating (3.3). The calculation of the requisite amplitudes b_{if} goes through as in § 4.1 as far as the formula (4.1), by selecting from the perturbation V' (3.5) just one multipole order λ.† The integral $S_{E\lambda,\mu}$ (4.2) in b_{if} now refers to a hyperbolic Rutherford orbit. For $|b_{if}|^2$ averaged over M_i and summed over M_f we have

$$\overline{|b_{if}|^2} = \left(\frac{4\pi Z_1 e^2}{\hbar}\right)^2 \frac{B(E\lambda)}{(2\lambda+1)^3} \sum_\mu |S_{E\lambda,\mu}|^2, \qquad \ldots\ldots(5.1)$$

where $B(E\lambda)$ is defined by (4.6). We gave an approximation to $\overline{|b_{if}|^2}$ in (4.5), but now we require it exactly in terms of the exact integrals $S_{E\lambda,\mu}$. These can be calculated by introducing a suitable parametric representation of the hyperbolic

† As explained in § 4.1, usually only one λ value need be considered, but if more than one is permitted by the selection rules, the respective cross sections $\sigma_{E\lambda}$ are additive (both differential and total cross sections).

orbit, with the result

$$S_{E\lambda,\mu} = v^{-1} a^{-\lambda} Y_{\lambda\mu}(\tfrac{1}{2}\pi, 0) I_{\lambda\mu}(\theta, \xi), \qquad \ldots\ldots(5.2)$$

where θ is the Rutherford scattering angle for the orbit, related to the impact parameter p by (4.19), ξ is the adiabatic parameter (4.12), and $I_{\lambda\mu}$ is a complicated dimensionless integral (Alder, Bohr *et al.* 1956, equation II.A.26). The result applies for a coordinate frame in which the orbit lies in the plane $\theta_p = \frac{1}{2}\pi$.

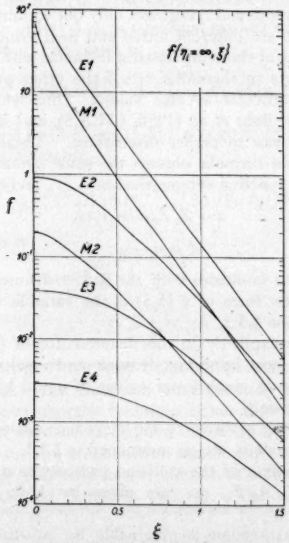


Figure 3. Total cross section functions $f_\lambda(\xi)$ in semi-classical approximation. (*After* Alder, Bohr *et al.* 1956.)

By making the various substitutions indicated above, we finally obtain the total cross section

$$\sigma_{E\lambda} = \left(\frac{Z_1 e^2}{\hbar v}\right)^2 \frac{B(E\lambda)}{a^{2\lambda-2}} f_{E\lambda}(\xi), \qquad \ldots\ldots(5.3)$$

where $f_{E\lambda}(\xi)$ is a complicated numerical function of ξ, involving integrals of the $I_{\lambda\mu}$'s over θ. Curves of $f_{E\lambda}(\xi)$ are given in figure 3. This result is similar in form to the previous estimate (4.10), (4.11), which gives a good approximation to (5.3) in the limit of small ξ, when $f_{E\lambda}$ tends to a constant (for $\lambda > 1$) and $f_{E1} \propto \ln(1/\xi)$. As expected, $f_{E\lambda}(\xi)$ falls off rapidly with increasing ξ (i.e. decreasing projectile energy), the perturbation becoming adiabatic.

It is a valuable simplicity in the process that the only non-elementary factor in the cross section, apart from the unknown quantity $B(E\lambda)$, is a calculable function

$f_{E\lambda}(\xi)$ of a single parameter ξ. $f_{E\lambda}(\xi)$ has been calculated numerically for $\lambda = 1-4$ by Alder and Winther (1954 b, 1956), and tables are also given in Alder, Bohr *et al.* (1956), tables II, 3, II, 4. In the dipole case only, an analytical formula can be obtained for $f_{E\lambda}(\xi)$ (Ter-Martirosyan 1952, Alder, Bohr *et al.* 1956, § II, E.5).

In applying the formulae of this method to actual experiments, one problem arises in giving a value to the velocity and energy of the projectile, in order to find v, a, ξ etc. The semi-classical treatment neglects the loss of energy of the projectile on its orbit due to excitation, and recognizes only one terminal velocity v or energy T, whereas in fact there are different initial and final velocities v_i, v_f. Does one obtain a closer agreement of the semi-classical formulae with actuality (or a quantal calculation) if one inserts in them for v and the other parameters a particular choice v_i, v_f or some particular average value? One average choice is indeed best, as shown by Alder, Bohr *et al.* (1956, § II.A.5), and Biedenharn and Thaler (1956), who give references to earlier discussions. These authors require that the classical cross section formula possess the same symmetry as the quantum mechanical ($\sigma = (v_f/v_i) \times$ function symmetrical in v_i, v_f), and this leads to the choice

(i) $$a = Z_1 Z_2 e^2/m_0 v_i v_f, \qquad \ldots\ldots(5.4)$$

(ii) $$\xi = \eta_f - \eta_i; \qquad \ldots\ldots(5.5)$$

it is readily seen that this coincides with the earlier definition of ξ, (4.12), when $\Delta E/T$ is small. The new form of ξ (5.5) is the variable which arises naturally in the quantum calculation § 5.2.

(iii) Where v occurs explicitly in the denominator of (5.3), v_i is to be used. The choice of ξ is the most important feature, and results in the semi-classical formulae remaining good to much lower values of η and higher values of ξ than would otherwise be the case.

A further important rule is always to use the reduced mass m_0, relative velocities of projectile and target nucleus, etc. as mentioned in § 3.1. Pre-digested formulae for the immediate evaluation of the collision parameters a, ξ, v_f etc. in terms of projectile energy, A_1, Z_1, A_2, Z_2, etc. are given by Alder, Bohr *et al.* (1956, § II.C.1).

The differential cross section is obtainable by substituting (5.1) in (3.11), employing also (5.2). The result is

$$\sigma_{E\lambda}(\theta)\, d\Omega = \left(\frac{Z_1 e^2}{\hbar v_i}\right)^2 \frac{B(E\lambda)}{a^{2\lambda-2}}\, df_{E\lambda}(\theta, \xi), \qquad \ldots\ldots(5.6)$$

where $df_{E\lambda}(\theta, \xi)$ is a function involving squares of $I_{\lambda\mu}(\theta, \xi)$'s, and

$$\int df_{E\lambda}(\theta, \xi) = f_{E\lambda}(\xi). \qquad \ldots\ldots(5.7)$$

Angular distribution curves for $\lambda = 1$–4 are shown in figure 2; tables of $df_{E\lambda}/d\Omega$ are given by Alder, Bohr *et al.* (1956, tables II.7 and II.8) for $\lambda = 1, 2$, and for $\lambda = 1$–4 by Alder and Winther (1956). These have been discussed in § 4.3.

5.2. *Quantal Calculation (method of Coulomb-distorted waves)*

According to the theory of § 3.3 the differential cross section is given by (3.15), using the potential V' given by (3.5) in the multipole expansion (2.9). The λ, μ

term in the matrix element $\langle \phi_f, \psi_f | V' | \phi_i, \psi_i \rangle$ factorizes into a nuclear matrix element proportional to

$$\left\langle \psi_f \left| \sum_k r_k^{\lambda} Y_{\lambda\mu}^*(\theta_k, \phi_k) \right| \psi_i \right\rangle$$

and a projectile matrix element proportional to

$$\langle \phi_f | r_p^{-(\lambda+1)} Y_{\lambda\mu}(\theta_p, \phi_p) | \phi_i \rangle.$$

We select one multipole order† λ, sum over final nuclear substates M_f and average over initial substates M_i. The result is

$$\sigma_{E\lambda}(\theta)\, d\Omega = \left(\frac{Z_1 e^2}{\hbar v_i}\right)^2 \frac{B(E\lambda)}{a^{2\lambda-2}}\, df_{E\lambda}(\theta, \eta_i, \xi), \qquad \ldots\ldots(5.8)$$

where a and ξ are defined by (5.4) and (5.5); η_i is given by (3.8) with v_i substituted for v, and $df_{E\lambda}$ (dimensionless) is

$$df_{E\lambda}(\theta, \eta_i, \xi) = \frac{4k_i k_f}{(2\lambda+1)^3}\, a^{2\lambda-2} \sum_{\mu} \left| \langle \phi_f | r_p^{-(\lambda+1)} Y_{\lambda\mu}(\theta_p, \phi_p) | \phi_i \rangle \right|^2 d\Omega. \qquad \ldots\ldots(5.9)$$

The total cross section is

$$\sigma_{E\lambda} = \left(\frac{Z_1 e^2}{\hbar v_i}\right)^2 \frac{B(E\lambda)}{a^{2\lambda-2}}\, f_{E\lambda}(\eta_i, \xi), \qquad \ldots\ldots(5.10)$$

where

$$f_{E\lambda}(\eta_i, \xi) = \int df_{E\lambda}(\theta, \eta_i, \xi).$$

The matrix elements in (5.9) between Coulomb wave functions of the projectile are difficult to evaluate. The Coulomb wave function $\phi(\mathbf{k}, \mathbf{r}_p)$ is expressible in terms of confluent hypergeometric functions either directly in closed form, or as a sum of partial waves, over orbital angular momentum l. However, we have in (5.8), (5.10) formulae for the cross section which differ formally from the semi-classical formulae (5.6), (5.3) only in the addition of a further parameter η_i in the arguments of the numerical functions $df_{E\lambda}, f_{E\lambda}$. The semi-classical f's are the limit of the quantal f's in the limit $\eta_i \to \infty$, ξ finite. In the dipole case, $df_{E1}(\theta, \eta_i, \xi)$ and $f_{E1}(\eta_i, \xi)$ can be evaluated in closed analytical form in terms of hypergeometric functions (Mullin and Guth 1951) by using the closed form of $\phi(\mathbf{k}, \mathbf{r}_p)$. In all other cases the partial wave expansion of ϕ must be employed, resulting in f's which are sums of quadratic terms in radial matrix elements $M_{l_i l_f}^{-\lambda-1}$:

$$M_{l_i l_f}^{-\lambda-1} = \frac{1}{k_i k_f} \int_0^{\infty} F_{l_f}(k_f r) r^{-\lambda-1} F_{l_i}(k_i r)\, dr, \qquad \ldots\ldots(5.11)$$

where F_l is the regular radial Coulomb wave function. Expressions for the radial integrals $M_{l_i l_f}^{-\lambda-1}$ in terms of generalized hypergeometric functions were obtained by Biedenharn, McHale and Thaler (1955) and Alder and Winther (1955 a, b). Numerical calculations of the total cross section for a range of ξ and η_i have been made by Alder and Winther (1955 c) for $\lambda = 2$, Biedenharn, Goldstein *et al.* (1956) for $\lambda = 2$, and Thaler *et al.* (1956) for $\lambda = 1$. Graphs and tables are also given by Alder, Bohr *et al.* (1956, table II.5). Biedenharn and Thaler (1956) have

† Again the cross sections for different λ are simply additive.

calculated the total cross section as a function of η_1 when $\xi = 0$ for $\lambda = 3, 4$. Very large numbers of partial waves l had to be taken into account. The major contributions (for $\lambda > 1$) come from $l \sim \eta$, for in the straight-orbit picture the most important impact parameters are $p \sim a$, which corresponds to angular momentum of the order $m_0 va \sim \hbar\eta$ (the last step follows from 3.10).

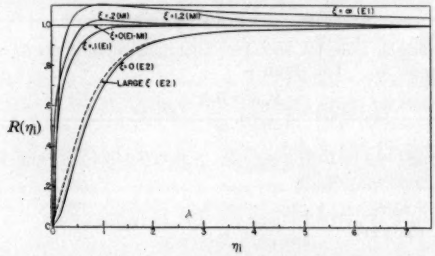


Figure 4. Ratio $R(\eta_1, \xi)$ of quantal total cross sections to semi-classical, for various multipoles and various specified values of ξ. (*After* Biedenharn and Thaler 1956.)

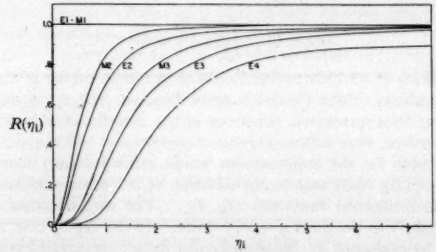


Figure 5. The ratio function $R(\eta_1, \xi)$ in the limit $\xi = 0$ for various multipoles. (*After* Biedenharn and Thaler 1956.)

It turns out that the quantal total cross section is very close to the classical limit $\eta_1 = \infty$ for values of η_1 down to approximately 1, as shown in figure 4, where the ratio $R(\eta_1, \xi)$ of the quantal cross section to the classical limit is plotted as a function of η_1 for fixed ξ values, in the cases $\lambda = 1, 2$. It is further noteworthy that R is only weakly dependent on ξ. Biedenharn and Thaler (1956) have therefore suggested that in cases where $f_{\mathrm{E}\lambda}(\eta_1, \xi)$ is not available, a good approximation may be obtained by correcting the semi-classical cross section by the ratio $R(\eta_1, 0)$ of quantal to classical, evaluated for the actual η_1, but for $\xi = 0$. Curves of these correction factors are shown in figure 5.

The shape of these curves can be understood from the approximate theory of § 4.1. According to this the ratio of the cross section in the case of Born approximation, which applies when $\eta_1 \ll 1$, to that given by the semi-classical

theory is of the order of $(ka)^{2\lambda-2}$, which equals $\eta^{2\lambda-2}$ (3.10). To represent a smooth transition between large and small η one may therefore estimate

$$R(\eta_1, 0) \sim \frac{1}{(1+\eta^{-2})^{\lambda-1}}, \qquad \ldots\ldots(5.12)$$

which does indeed represent the trends in figure 5.

Numerical calculations of the differential cross section have not been made, but the formulae are given by Alder, Bohr *et al.* (1956, § II.B.3).

5.3. *Limitations of Present Theory*

The limitations of the above treatment have been discussed by Biedenharn, McHale and Thaler (1955) and Alder and Winther (1956, § II.C.5). The principal effects which have been neglected in the calculations are as follows:

5.3.1. *Electromagnetic corrections* (§ 2).

The condition for retardation effects to be negligible is (2.8) for which, except in the dipole case, we may put $r_{p,\max} \sim a$. It is then readily found that (2.8) is equivalent to

$$\xi v/c \ll 1. \qquad \ldots\ldots(5.13)$$

v/c is always rather small ($\sim 0{\cdot}15$ for 10 MeV protons), and so the condition is well satisfied, particularly as ξ is usually kept fairly small. In the dipole case we write $r_{p,\max} = \hat{p}_{\max}$ (4.9) and then condition (2.8) becomes

$$v/c \ll 1. \qquad \ldots\ldots(5.14)$$

A further electromagnetic correction mentioned in § 2 is the contribution to electric transitions from $V_A^{(1)}$ but this interaction was shown to be small compared with V.

5.3.2. *Finite nuclear size.*

The use of waves ϕ distorted by the point charge interaction V_0, and the use of the multipole interaction (2.9) for all values of r_p are only correct if the nucleus is negligibly small, or if Coulomb repulsion completely prevents penetration of the nucleus. Possible, though very difficult, improvements would be the use of an optical-model potential in place of V_0 to determine the distorted waves, and the use of the correct potential V of (2.9) and (2.10). The error involved in these effects is indicated by the probability of barrier penetration, which is governed by a factor approximately

$$\exp\left\{-2\pi\eta\left[1-\frac{4}{\pi}\left(\frac{T}{E_B}\right)^{1/2}\right]\right\}, \qquad \ldots\ldots(5.15)$$

provided that T is appreciably less than the Coulomb barrier E_B (cf. Bethe 1937, p. 163). This shows that provided η is large and the energy is well below barrier height, barrier penetration will indeed be small. However, for smaller η this is no longer true, and the corrections here envisaged may become important. We see that the principal criterion for smallness of these corrections is the largeness of η—the same criterion as for the validity of the semi-classical theory of § 5.1 relative to the quantal theory of § 5.2. This suggests that the replacement of

the semi-classical theory by the quantum theory as in § 5.2 is called for only in conditions which would require the additional corrections for finite nuclear size to be made simultaneously.

We have been considering the effect of finite nuclear size only on the Coulomb excitation process, but it also causes competing nuclear reactions. Expression (5.15) for the barrier penetration indicates that, provided η is large and the energy is well below barrier height, nuclear reactions will be suppressed below Coulomb excitation, which is subject to depression only by the weaker factor $\exp(-2\pi\xi)$ (cf. 4.12).

5.3.3. *Electronic screening.*

The effect of this will be small provided the distance of closest approach is much smaller than the radius of the atomic shell orbit r_K. An elementary calculation yields

$$\frac{a}{r_K} \sim \frac{\eta^2 m_e}{Z_1 m_0}, \qquad \ldots\ldots (5.16)$$

and so this is small except for exceedingly large η, i.e. in all practical circumstances.

5.3.4. *Higher order effects.*

Both the classical and quantal calculations treat the excitation by first-order perturbation theory, but the effects of higher order have been discussed by Breit and Lazarus (1935), Breit, Gluckstern and Russell (1956, 1957), and Alder, Bohr *et al.* (1956, § II.D) in the semi-classical approximation.

Obviously the first-order perturbation treatment is worse the larger the excitation cross section. Its validity can be gauged by the smallness of the transition amplitudes b_{if} (3.1): to estimate these we may take the square root of our formula (4.5), the largest value of b_{if} occurring on the orbit which makes $p = 2a$. Then it turns out that under ordinary experimental conditions b_{if} is reasonably small. Thus, to take an extreme example, 7MeV alphas on $^{158}_{64}$Gd, for which $B(E2) = 10^{-47}$ cm^4, yield a maximum $b_{if} \sim 0{\cdot}12$, and the same particles on $^{23}_{11}$Na, for which $B(E2) = 1{\cdot}3 \times 10^{-50}$ cm^4, yield $b_{if} \sim 0{\cdot}15$. Consequently the E2 cross section calculations relating to existing data are reliable, but might become less so under conditions producing higher cross sections. Thus if we use different projectiles, making their energies all the same fraction of their respective Coulomb barriers (cf. § 7.2), we see from (4.5) that $b_{if} \propto (Z_1 A_1)^{1/3}$, so that the theory is worse for heavier projectiles.

In second-order perturbation theory, the transition $i \to f$ proceeds through intermediate nuclear states $i \to n \to f$. In a crude estimate this replaces the amplitude b_{if}, as given by (4.5), by something like $\Sigma_n b_{in} b_{nf}$. This may be very important for a transition which proceeds in first order by a large multipole λ, e.g. $I_i = 0 \to I_f = 4$ requiring $\lambda = 4$. It can go in second order instead by two virtual quadrupole transitions, via an intermediate state for which $I_n = 2$. The largeness of E2 b's relative to E4 ones may make the second-order process quite favourable. Calculation (assuming small ξ) shows that on the single-particle model (§ 4.2) the ratio of the second-order cross section to the first order in such a transition is about $15(Z_1 e^2/\hbar v)^2$, which could be greater than unity for α and heavier particles; and the ratio may be much greater if the quadrupole transitions

are collective. If $\xi = \eta_f - \eta_i$ becomes large, the cross sections for second and first order transitions are expected to fall off in a similar fashion, irrespective of the intermediate state energy.

One second-order effect in quadrupole transitions, obtained by putting n = i or n = f, is the introduction into the transition amplitude of terms proportional to the static quadrupole moment of the initial state i or the final state f.

5.4. *J.B.W.K. Method*: *Bridge between Semi-Classical and Quantal Theories*

It was an interesting theoretical problem to demonstrate formally how the quantal calculation (§ 5.2) goes over into the semi-classical (§ 5.1) in the limit of large η, and also to investigate the degree of goodness of the semi-classical limit. The link between the methods is the application of the Jeffries–Brillouin–Wentzel–Kramers approximation to the quantal calculation. This has been performed by Alder and Winther (1954 a), Breit and Daitch (1954), Daitch *et al.* (1954), Breit and Daitch (1955), Lazarus and Sack (1955), Alder and Winther (1955 b), Benedict *et al.* (1956), Gluckstern *et al.* (1956) and Alder, Bohr *et al.* (1956, § II.B.6). On employing the lowest order J.B.W.K. approximation for the radial wave functions in the matrix element $M_{l_i l_f}{}^{-\lambda-1}$ (5.11) and treating $(k_i - k_f)$ and $(l_i - l_f)$ as small quantities, one obtains an expression for $M_{l_i l_f}{}^{-\lambda-1}$ closely related to the semi-classical orbital integral $S_{\mathrm{E}\lambda,\,(l_f - l_i)}$ (4.2), (provided that for the latter axes are chosen so that the plane of scattering is the plane $\theta_p = \frac{1}{2}\pi$). The precise relation, in terms of the integrals $I_{\lambda\mu}$ introduced in (5.2), turns out to be

$$M_{l_i l_f}{}^{-\lambda-1} = \frac{1}{4a^{\lambda}\, k_i\, k_f}\, I_{\lambda\,(l_f - l_i)}\,(\theta, \xi). \qquad \ldots\ldots(5.17)$$

The semi-classical orbit here is identified by the Rutherford scattering angle θ as given by (4.19), taking the impact-parameter p related to the quantal orbital angular momentum by

$$m_0 v p = \hbar[l(l+1)]^{1/2}, \qquad \ldots\ldots(5.18)$$

with

$$l = \tfrac{1}{2}(l_i + l_f)\,; \qquad \ldots\ldots(5.19)$$

this is equivalent to the correspondence

$$\eta \cot \tfrac{1}{2}\theta = [l(l+1)]^{1/2} \qquad \ldots\ldots(5.20)$$

in which we take

$$\eta = (\eta_i \eta_f)^{1/2}. \qquad \ldots\ldots(5.21)$$

Application of the usual conditions for the J.B.W.K. method shows that (5.17) is valid provided *either* η *or* the l's tend to infinity, $l_f - l_i$ be not large (in fact we have $|l_f - l_i| \leqslant \lambda$), $\Delta E \ll T$ and ξ remain finite. Moreover the investigations mentioned above have established that the approximate correspondence (5.17) is unexpectedly accurate even for moderate values of η and l, provided one uses the 'symmetrized' values of the classical parameters as in (5.19)–(5.21), (5.4), (5.5), etc.

To obtain the total cross section, (5.17) is substituted in the series which expresses the quantal $\sigma_{\mathrm{E}\lambda}$ in terms of the M's, the coefficients in this series are simplified for the limit of large l, and the sum over l is replaced by an integral over p (cf. (5.18)): this finally yields the result of the semi-classical method (5.3).

A similar but more complicated calculation shows that the quantal formula for the differential cross section also reduces to the semi-classical (5.6), with the aid of (5.17).

The relation (5.17) was derived differently by Biedenharn *et al.* (1955) under the same limiting conditions on the parameters, by a direct mathematical transformation of the limiting value of the integral on the left-hand side into the one on the right. The goodness of the semi-classical formula for the total cross section was already evidenced in figures 4 and 5. The differential cross section must be less accurately represented by the semi-classical approximation, since the quantal formula (unlike that for the total cross section) contains interference terms between products of pairs of M's together with phase factors, and the sum of these terms is expected to show poor convergence to the classical limit.

5.5. *Angular Distribution of De-excitation γ-Rays*

The excited state produced in the target nucleus will decay with emission of a γ-ray (plus internal-conversion electrons), and the γ-ray will have an angular distribution with azimuthal symmetry about the direction of the incident projectile beam. The γ-ray will further have a more complicated correlation with the direction of the inelastically scattered particles, but this has been little investigated. The polarization of the de-excitation γ-rays can also be considered.

Any de-excitation or other phenomena occurring after excitation are determined by the density-matrix $\rho_{M_f,M_f'}$ of the excited state. This is given on the semi-classical theory in terms of the transition amplitudes b_{if} (3.1):

$$\rho_{M_f,M_f'} = \frac{1}{(2I_i+1)} \sum_{M_i} \int b_{M_f,M_i} b_{M_f,M_i'}^{*}\, p\, dp\, d\phi. \qquad \ldots\ldots(5.22)$$

Here, integration over orbit parameters p and ϕ corresponds to non-observation of the inelastically scattered particle. In the quantal theory, the density matrix is given in terms of the matrix elements (3.15):

$$\rho_{M_f,M_f'} = \frac{1}{(2I_i+1)} \sum_{M_i} \int \langle \vec{\phi}_f, \psi_{M_f} | V' | \phi_i, \psi_{M_i} \rangle \langle \vec{\phi}_f, \psi_{M_f'} | V' | \phi_i, \psi_{M_i} \rangle^{*} \sin\theta\, d\theta\, d\phi. \qquad \ldots\ldots(5.23)$$

Again we have integrated over directions (θ, ϕ) of inelastic scattering. The physical effects depend directly on statistical tensors formed from the density matrix elements with vector addition coefficients (see for example Devons and Goldfarb 1957, Biedenharn and Rose 1953)

$$\rho_{k,\kappa} = \sum_{M_f,M_f'} (-1)^{I_f - M_f'} (I_f I_f M_f - M_f' | I_f I_f k\kappa)\, \rho_{M_f,M_f'}. \qquad \ldots\ldots(5.24)$$

Owing to parity conservation k here takes only even integral values, not larger than $2I_f$; and the azimuthal symmetry means that only the $\kappa = 0$ tensor components are non-zero (assuming the principal coordinate axis to be in the direction of incidence). Similar density matrices and statistical tensors would arise if the nuclear transition $I_i \to I_f$ were induced by another mechanism specified by the same symmetry axis, such as absorption of an unpolarized γ-ray incident along the $\mathbf{k}_i$ direction. It turns out that (provided only a single multipole λ is involved)

the statistical tensors factorize into a ' nuclear factor ' which, while depending on I_i, I_f, λ, k, is the same for different mechanisms (apart from a normalization factor), and a tensor ' radiation parameter ' $c_{k0}(\lambda\lambda)$ which is characteristic of the mechanism. It is thus convenient to compare the Coulomb excitation with the corresponding γ excitation of order Eλ, and write

$$\frac{\rho_{k0}(\text{Coulomb})}{\rho_{k0}(\gamma)} = \frac{c_{k0}(\lambda\lambda)(\text{Coulomb})}{c_{k0}(\lambda\lambda)(\gamma)} = a_k^{\mathrm{E}\lambda}. \qquad \ldots\ldots(5.25)$$

The tensors here are normalized so that $a_0^{\mathrm{E}\lambda} = 1$. For specified λ, k cannot exceed 2λ. Actually, the comparison has conventionally been made with a hypothetical *γ-emission* from I_i to I_f instead of a *γ-absorption*. The c's for γ-emission are well known (Devons and Goldfarb 1957, § 14)

$$c_{k0}(\lambda\lambda)(\gamma) = \frac{(2\lambda+1)}{4\pi}(-1)^{\lambda-1}(\lambda\lambda 1 - 1|\lambda\lambda k0). \qquad \ldots\ldots(5.26)$$

For Coulomb excitation one finds for the semi-classical case from (5.22), before normalization,

$$c_{k0}(\lambda\lambda)(\text{Coulomb}) = \sum_{\mu}\int(-1)^{\lambda-\mu}(\lambda\lambda\mu-\mu|\lambda\lambda k0)\, S_{\mathrm{E}\lambda,\mu}\, S_{\mathrm{E}\lambda,\mu}^{*}\, p\, dp\, d\phi, \qquad \ldots\ldots(5.27)$$

and for the quantal case from (5.23)

$$c_{k0}(\lambda\lambda)(\text{Coulomb}) = \sum_{\mu}\int(-1)^{\lambda-\mu}(\lambda\lambda\mu-\mu|\lambda\lambda k0) \langle\phi_f|r_p^{-(\lambda+1)}\mathrm{Y}_{\lambda\mu}(\theta_p,\phi_p)|\phi_i\rangle\langle\phi_f|r_p^{-(\lambda+1)}\mathrm{Y}_{\lambda\mu}(\theta_p,\phi_p)|\phi_i\rangle^{*}\sin\theta\, d\theta\, d\phi. \qquad \ldots\ldots(5.28)$$

The angular distribution of a γ-ray emitted in de-excitation from the state f to a state ff may now be obtained by comparison with a hypothetical radioactive γ–γ cascade $I_i \to I_f \to I_{ff}$, which will have a correlation function

$$W(\theta) = \sum_k A_k P_k(\cos\theta), \qquad \ldots\ldots(5.29)$$

where θ is the angle between the directions of the successive γ's. Now the coefficient A_k is proportional to ρ_{k0}, and therefore we obtain the angular distribution for de-excitation after the Coulomb process by simply inserting the factors $a_k^{\mathrm{E}\lambda}$:

$$W(\theta\gamma) = \sum_k A_k a_k^{\mathrm{E}\lambda} P_k(\cos\theta_\gamma), \qquad \ldots\ldots(5.30)$$

where θ_γ is the angle between the γ-ray and the incident beam. The coefficients A_k for γ–γ cascades are obtainable from the literature (Biedenharn and Rose 1953, Ferentz and Rosenzweig 1954, Alder, Bohr *et al.* 1956, Table II, 11). The first ' hypothetical ' ray $I_i \to I_f$ is to be taken as pure Eλ, but the second ' actual ' γ-ray $I_f \to I_{ff}$ must allow for any actual decay E/M admixture.

The calculation of $a_k^{\mathrm{E}\lambda}$ is very similar in character to the calculation of the total cross section, as may be seen by comparing (5.27) with (5.1) or (5.28) with

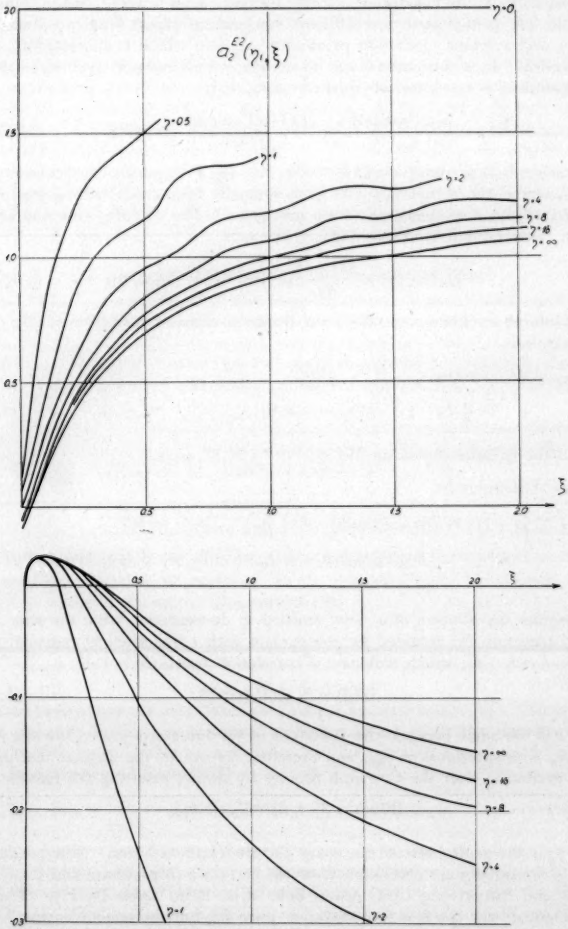


Figure 6. Decay γ-ray angular distribution coefficients a_2 and a_4 for E2 excitations, as functions of ξ for various values of η_1. (*After* Alder and Winther 1955 b.)

(5.9). Thus $a_k^{E\lambda}$, in the semi-classical case, is a function only of ξ and, in the quantal case, only of η_i and ξ (apart from affixes k, λ).

Formulae, curves and tables for the a's in the cases $\lambda = 1, 2$ were given on the semi-classical theory by Alder and Winther (1953, 1954 b). Quantal calculations for $\lambda = 2$ were made by Biedenharn and Class (1955 a). However, all these calculations were bedevilled by numerical errors, which were sorted out in recalculations by Biedenharn, McHale and Thaler (1955), Breit, Ebel and Benedict (1955), Goldstein *et al.* (1955), Benedict and Tice (1955), Biedenharn and Class (1955 b), Benedict (1956) and Breit, Ebel and Russell (1956). Correct tables and curves for $\lambda = 1, 2$ are given by Alder and Winther (1955 b, c), Alder, Bohr *et al.* (1956, § II.C.4), Biedenharn, Goldstein *et al.* (1956), and Thaler *et al.* (1956). Figure 6 shows curves of $a_2^{E2}(\eta_i, \xi)$, $a_4^{E2}(\eta_i, \xi)$. The semi-classical limit ($\eta_i = \infty$) is less good as an approximation to the quantal curves of finite η_i than in the analogous case of the total cross section (cf. figure 4). This is because the formulae for the γ angular coefficients contain interference terms which are absent from the total cross section.

The same limitations on the validity of the calculations apply as in § 5.3. To calculate other de-excitation effects, such as the polarization of the γ-rays, or possible disorienting effects of extra-nuclear fields, it is possible again to start with the formula referring to a comparison process initiated by a hypothetical $E\lambda$ γ-emission in place of the Coulomb excitation, and then to apply the correction coefficient $a_k^{E\lambda}$ to each term in the comparison formula containing the tensor suffix k, as in (5.29) and (5.30).

§ 6. Excitation by Electrons

6.1 *Electrostatic Interaction in Born Approximation*

The simplest approximation, following Schiff (1954), is to use the electrostatic potential (2.1) in the Born approximation (3.13). The integral over the electron coordinates in the matrix element can readily be performed, and in the extreme relativistic limit we find for the differential cross section

$$\sigma(\theta) = 4\left(\frac{e^2}{\hbar c}\right)^2 \frac{k_f^2 \cos^2 \tfrac{1}{2}\theta}{K^4} \left| \left\langle \psi_f \middle| \sum_k \exp(i\mathbf{K}.\mathbf{r}_k) \middle| \psi_i \right\rangle \right|^2, \qquad \ldots\ldots(6.1)$$

where

$$\mathbf{K} = \mathbf{k}_i - \mathbf{k}_f. \qquad \ldots\ldots(6.2)$$

Using (4.22) for K and assuming $\Delta E \ll T$, we can write the result (except at very small scattering angles):

$$\sigma(\theta) = \left(\frac{Z_2 e^2}{\hbar c}\right)^2 \frac{\cos^2 \tfrac{1}{2}\theta}{4k_i^2 \sin^4 \tfrac{1}{2}\theta} |F(I_i \to I_f, K)|^2, \qquad \ldots\ldots(6.3)$$

where

$$F(I_i \to I_f, K) = \frac{1}{Z_2} \left\langle \psi_f \middle| \sum_k \exp(i\mathbf{K}.\mathbf{r}_k) \middle| \psi_i \right\rangle. \qquad \ldots\ldots(6.4)$$

The expression (6.3) without its last factor is the differential cross section for elastic scattering by a point charge $Z_2 e$, in relativistic Born approximation, and the last factor can be regarded as a 'form factor', which depends on the projectile parameters only through K. The elastic scattering by the nucleus could also be obtained from (6.3) and (6.4) by putting $\psi_f = \psi_i$. The plane waves $\exp(i\mathbf{K}.\mathbf{r}_k)$ in (6.4) are now expanded in spherical harmonics P_λ. The transition $I_i \to I_f$

selects just the multipole orders λ permitted in Eλ transitions (§ 2) (but including $\lambda = 0$), and as usual we retain only one of these (noting that the cross section ought really to be summed over λ). We find after summing and averaging over M_f, M_i:

$$|F_{E\lambda}(I_i \to I_f, K)|^2 = \frac{4\pi K^{2\lambda}}{Z_2^2[(2\lambda+1)!!]^2} \cdot B(C\lambda, K), \qquad \ldots\ldots(6.5)$$

where $B(C\lambda, K)$ is a nuclear transition probability defined in the same way as the radiative transition probability $B(E\lambda)$ (4.6), except that in the latter formula r_k^λ is replaced by the substitution

$$r_k^\lambda \to \frac{(2\lambda+1)!!}{K^\lambda} j_\lambda(Kr_k). \qquad \ldots\ldots(6.6)$$

This notation is used because in the limit of small KR_0 (6.6) becomes an identity, and so

$$B(C\lambda, 0) = B(E\lambda). \qquad \ldots\ldots(6.7)$$

In order to estimate the behaviour, except in the case $\lambda = 0$, we may approximate $B(C\lambda, K)$ by $B(E\lambda)$, provided KR_0 is not large. The total cross section obtained in this approximation by integration of the differential has been given in (4.13), (4.14). We can also understand how the cross section for larger KR_0 falls progressively below those estimates owing to diffraction effects, for in (6.6) the right-hand side becomes smaller than the left-hand side, and so $B(C\lambda, K)$ drops below $B(E\lambda)$. For the angular distribution $\sigma(\theta)$ we obtain from the above formulae a factor $K^{2\lambda-4} \propto (\sin\frac{1}{2}\theta)^{2\lambda-4}$ as in § 4.3, but there are additional factors: (i) a factor $\cos^2\frac{1}{2}\theta$, which is a relativistic effect of the electron spin, (ii) a dependence of $B(C\lambda, K)$ on K and hence on θ. The last factor is a diffraction effect, as (6.6) shows, and its influence on the angular distribution constitutes the 'probing' of the nucleus by the electrons.

The monopole case $\lambda = 0$ has also to be taken into account. Using the long wavelength limit (6.7), we see from (4.6) that $B(E0)$ is $Z_2^2/4\pi$ for elastic scattering and zero for inelastic, owing to orthogonality of the initial and final states. Hence, in the limit, for elastic scattering the monopole form factor (6.5) is unity, as expected for long wavelength; and there is no inelastic scattering. Finite monopole inelastic scattering is only obtained by going to the second term in the power series for $j_\lambda(Kr_k)$ in (6.6):

$$j_\lambda(x) = \frac{x^\lambda}{(2\lambda+1)!!}\left[1 - \frac{x^2}{2(2\lambda+3)} + \ldots\right] \qquad \ldots\ldots(6.8)$$

This is substituted in the formula for $B(C\lambda, K)$, and then (6.5) yields

$$|F_{E0}(I_i \to I_f, K)|^2 = \frac{K^4}{36Z_2^2}\left|\left\langle I_f, M_f \middle| \sum_k r_k^2 \middle| I_i, M_i \right\rangle\right|^2. \qquad \ldots\ldots(6.9)$$

This exhibits the same energy and angle dependence as the quadrupole form factor given by (6.5) in the long wavelength limit, (6.7). The matrix element squared in (6.9) may be estimated on the single-particle model as in (4.15), and this yields $(3/5)^2 R_0^4$. As a result this model predicts that any monopole differential cross section just about equals any quadrupole under the same experimental conditions. Monopole matrix elements of the order suggested by the single particle model are known in light elements.

To facilitate discussion of the variation of $|F_{E\lambda}(K)|^2$ with K, we can define a λ-order radial transition density for protons in the nucleus $\rho_\lambda(I_1 \to I_f, r)$ by

$$\int \psi_{M_f}^*(\mathbf{r}_1, \sigma_1, \mathbf{r}_2, \sigma_2 \ldots)\, Y_{\lambda\mu}^*(\theta_1, \phi_1)\, \psi_{M_i}(\mathbf{r}_1, \sigma_1, \mathbf{r}_2, \sigma_2 \ldots)\, d\Omega_1\, d\sigma_1\, d\mathbf{r}_2\, d\sigma_2 \ldots$$
$$= (4\pi)^{1/2} (I_f\, \lambda M_f\, \mu \,|\, I_i\, \lambda I_i\, M_i)\, \rho_\lambda(I_1 \to I_f, r_1), \qquad \ldots\ldots(6.10)$$

the integral being taken over all nuclear coordinates (including spins σ) except r_1, which refers to a proton. ρ_λ is a generalization of the ordinary density of particles in the nucleus. Then (6.4) yields for $|F(I_i \to I_f, K)|^2$ (picking out only the λ term, and summing and averaging over M_f, M_i)

$$|F_{E\lambda}(I_1 \to I_f, K)|^2 = \left[4\pi \int_0^\infty \rho_\lambda(I_1 \to I_f, r) j_\lambda(Kr)\, r^2\, dr\right]^2. \qquad \ldots\ldots(6.11)$$

Thus the inelastic form factor measures a Fourier transform of the transition density, as the elastic does of the static density (cf. (6.16) below). We have a relation between ρ_λ and $B(E\lambda)$:

for $\lambda \neq 0$

$$B(E\lambda) = 4\pi \left[Z_2 \int_0^\infty \rho_\lambda(I_i \to I_f, r)\, r^{\lambda+2}\, dr\right]^2, \qquad \ldots\ldots(6.12)$$

or for monopole transitions

$$\left\langle I_i, M_i \middle| \sum_k r_k^2 \middle| I_i, M_i \right\rangle = 4\pi Z_2 \int_0^\infty \rho_0(I_i \to I_i, r) r^4\, dr. \qquad \ldots\ldots(6.13)$$

6.2. *Sum of Cross Sections over States*

While we have discussed the excitation of a particular final level, estimates have also been made of the differential cross section summed over final levels (Gatto 1953, 1955, Smith 1954, Lewis 1956, Schiff 1957). According to (6.3), we require the sum of the form factors over final states, though this is only approximately true, since (6.3) is an approximation to (6.2) valid only for $\Delta E \ll T$. However, with this simplification we get from (6.4), using the closure property of the wave functions ψ_f:

$$\sum_f |F(I_i \to I_f, K)|^2 = \frac{1}{Z_2^2} \left\langle \psi_i \middle| \sum_{k,k'} \exp\{i\mathbf{K}.(\mathbf{r}_k - \mathbf{r}_{k'})\} \middle| \psi_i \right\rangle, \qquad \ldots\ldots(6.14)$$

and by using the antisymmetry of the wave function finally

$$\sum_f |F(I_i \to I_f, K)|^2 = \frac{1}{Z_2} + \left(1 - \frac{1}{Z_2}\right) \langle \psi_i | \exp\{i\mathbf{K}.(\mathbf{r}_1 - \mathbf{r}_2)\} | \psi_i \rangle. \qquad \ldots\ldots(6.15)$$

This may be split up into its elastic and inelastic parts, and we can conveniently write

$$|F|_{el}^2 = \left[\int \rho_i(\mathbf{r}) \exp(i\mathbf{K}.\mathbf{r}) d\mathbf{r}\right]^2, \qquad \ldots\ldots(6.16)$$

and

$$|F|_{inel}^2 = \frac{1}{Z_2}\left\{1 - \left[\int \rho_i(\mathbf{r}) \exp(i\mathbf{K}.\mathbf{r})\, d\mathbf{r}\right]^2\right\}$$
$$- \left(1 - \frac{1}{Z_2}\right) \iint \rho_c(\mathbf{r}_1, \mathbf{r}_2) \exp\{i\mathbf{K}.(\mathbf{r}_1 - \mathbf{r}_2)\} d\mathbf{r}_1\, d\mathbf{r}_2, \qquad \ldots\ldots(6.17)$$

where $\rho_1(\mathbf{r})$ is the one-proton density function in the ground state defined by

$$\rho_1(\mathbf{r}_1) = \int |\psi_1(\mathbf{r}_1, \sigma_1, \mathbf{r}_2, \sigma_2 \ldots)|^2 \, d\sigma_1 \, d\mathbf{r}_2 \, d\sigma_2 \ldots, \qquad \ldots\ldots (6.18)$$

and $\rho_c(\mathbf{r}_1, \mathbf{r}_2)$ is a correlation or fluctuation density between two protons in the ground state, defined by

$$\rho_c(\mathbf{r}_1, \mathbf{r}_2) = \rho_1(\mathbf{r}_1)\rho_1(\mathbf{r}_2) - \int |\psi_1(\mathbf{r}_1, \sigma_1, \mathbf{r}_2, \sigma_2 \ldots)|^2 \, d\sigma_1 \, d\sigma_2 \, d\mathbf{r}_3 \, d\sigma_3 \ldots, \qquad \ldots\ldots (6.19)$$

and satisfying $$\int \rho_c(\mathbf{r}_1, \mathbf{r}_2) \, d\mathbf{r}_1 \, d\mathbf{r}_2 = 0.$$

Thus the elastic scattering at angle θ ('coherent scattering') is determined by a Fourier component of the ordinary proton density $\rho_1(\mathbf{r})$, and the summed inelastic ('incoherent scattering') is determined largely by a Fourier component of the correlation density $\rho_c(\mathbf{r}_1, \mathbf{r}_2)$.

Let us consider the order of magnitude of the form factors as functions of K according to these formulae. The entire form factor (6.14) is evidently always less than unity. When $K \ll R_0^{-1}$, elastic scattering, equal to point-nucleus scattering, will predominate. For values of $K \sim R_0^{-1}$, the cross section will generally be shared primarily between elastic scattering and excitation of collective levels, and the magnitude of both these effects will fall off in a similar way with increasing K, as governed by the Fourier components in (6.16) and (6.17). For $K \gg R_0^{-1}$ (projectile energies of several hundred MeV, and large angle scattering), excitation will take place mainly to levels of single-proton excitation or emission, which are favoured when the wavelength is of the order of the mean inter-nucleon separation. The elastic scattering is then very small, and both terms in the inelastic (6.17) will be of order $1/Z_2$. At still higher K, we shall simply have $|F|_{\text{inel}}^2 = 1/Z_2$, and so the cross section is equal to the Coulomb scattering from Z_2 independent point charges e. The behaviour with increasing K determines either the energy variation of cross section at a fixed angle, or the angular distribution at a fixed high energy.

Detailed calculations of the coherent and incoherent form factors have been made for specific nuclear models by Gatto (1953, 1955) and Smith (1954).

6.3. *Complete Electromagnetic Interaction*

The electromagnetic interaction is treated properly by using the sum of the interactions V and $V_A^{(1)}$ (2.1)–(2.3), with the retardation inserted as in (2.6). These interactions can be substituted in the Born formula for the cross section (3.13), in place of the electrostatic interaction used in § 6.1. This procedure is equivalent to the use of the Møller potentials by Schiff (1954), (who gives references to earlier work on these lines) and to another method adopted by Alder, Bohr *et al.* (1956, § II.E.3). These authors derive detailed formulae for the resulting differential cross sections for electric and magnetic multipole transitions. Here we consider only some of the features of the result.

The correct treatment really involves two separate corrections to the electrostatic approximation: (i) the effect of electromagnetic retardation, as represented

by the replacement (2.6), (ii) the addition of the vector-potential interaction $V_A^{(1)}$ to the scalar V. The effects of these corrections can be studied separately.

(i) *Retardation effects.* The cross sections obtained using the retarded Green's function (2.6) reduce to those using the unretarded, when $K \gg \kappa$; this agrees with the criterion (2.8) if we put $r_{p,\max} \sim 1/K$, a reasonable uncertainty relationship (cf. § 4.3). If we assume $k \gg \kappa$, which means $T \gg \Delta E$, then it follows, with the aid of (4.22), that retardation is only significant for small-angle scattering, such that

$$\theta \lesssim \Delta E/T. \qquad \ldots\ldots(6.20)$$

We note that this is the same region of small angles for which the cross section formula (6.3) used previously is not a good approximation to (6.1). Therefore for large angle scattering of high energy electrons we may use the quasi-static interactions V and $V_A^{(1)}$ of (2.1)–(2.3) as they stand without retardation.

(ii) *Vector potential interaction.* The addition of $V_A^{(1)}$ to the electrostatic theory of § 6.1 has the effect of adding to the cross section (6.3) for electric transitions a term similar in form to this, but obtainable from it by the following two alterations: firstly the 'spin' factor $\cos^2 \frac{1}{2}\theta$ in (6.3) is to be replaced by $\frac{1}{2}(1+\sin^2 \frac{1}{2}\theta)$; secondly the λ-order form factor is given not by (6.5) but by

$$|F_{\mathrm{E}\lambda(\mathrm{A})}(I_i \to I_f, K)|^2 = \frac{4\pi(1+1/\lambda)K^{2\lambda}}{Z_2^2[(2\lambda+1)!!]^2} \cdot B(\mathrm{E}\lambda, K), \qquad \ldots\ldots(6.21)$$

(no monopole)

where $B(\mathrm{E}\lambda, K)$ is defined in the same way as $B(\mathrm{E}\lambda)$ (4.6), except that in the latter the operator is replaced by the substitution (valid for large scattering angles only)

$$\sum_k r_k^{\lambda} \mathrm{Y}_{\lambda\mu}(\theta_k, \phi_k) \to \frac{i(2\lambda+1)!!}{ec(\lambda+1)K^{\lambda-1}} \int \mathbf{j}_n \cdot \mathbf{r} j_\lambda(Kr) \mathrm{Y}_{\lambda\mu}(\theta, \phi) d\tau. \qquad \ldots\ldots(6.22)$$

Both these operators have the same selection rules. The new matrix element, depending on the nuclear current density $\mathbf{j}_n$, contains an implicit factor v_k/c in comparison with the electrostatic matrix elements, and so the additional cross section can be expected to be rather small in comparison with the electrostatic (6.3), (6.4). It will, however, grow in importance as KR_0 increases.

We may conclude with Schiff that for large angle scattering of high energy electrons, the electrostatic approximation of § 6.1 is good. For small angles, the exact formulae of the references quoted above would have to be used.

6.4. *Further Refinements*

For heavy nuclei it would be necessary to use electron wave functions distorted by the static charge distribution of the nucleus, on the lines of the method of Coulomb distorted waves for nuclear projectiles § 3.3, but this would be very difficult. Schiff (1957) has developed an approximation resembling B.W.K. which would be usable when $kR_0 \gg 1$.

Higher order radiative corrections to scattering, in particular the simultaneous emission of soft photons, have been considered in elastic scattering of electrons (Schwinger 1949 a, b, Suura 1955) and are of some importance: the inelastic scattering will be similarly affected.

PART II. EXPERIMENT

§7. Excitation by Nuclear Projectiles

7.1. *Experimental Methods*

Most excitations have been produced with beams from electrostatic generators, furnishing protons with energies of up to about 5 MeV or α-particles (or He^+ ions) of up to about 7 MeV. The excitations have been observed mainly by means of the ensuing de-excitation radiations, either γ-rays or internal conversion electrons. In the case of γ-rays, both angular distributions (e.g. McGowan and Stelson 1957) and cross sections integrated over angle (e.g. Temmer and Heydenburg 1956 a, Heydenburg and Temmer 1956 b) have been determined ; in the case of conversion electrons, only the integrated cross section (e.g. Huus, Bjerregaard and Elbek 1956). Some measurements of γ-ray polarization have also been made (Stelson and McGowan 1956). The variation of cross section with energy has usually been measured (figures 7 and 8). Frequently only one or two excited states, with energies ranging from 30 kev to 1·5 MeV, are observed, occasionally as many as three or four.

The γ-rays have normally been detected with scintillation counters, and the conversion electrons with the wedge-gap type β-spectrometer. Conversion electrons have also been observed indirectly by detection of the succeeding X-rays (Heydenburg and Temmer 1956 b, Jones and Phillips 1957). Observation of conversion electrons is most favourable in the region of heavy nuclei and low energy levels ($\lesssim$ 200 kev), where the conversion coefficients are greatest. Better energy resolution is obtainable with conversion electrons than with γ-rays. Thick targets (i.e. ones stopping the charged projectiles) have been used more often than thin ones in order to enhance the intensity for γ-rays ; but with conversion electrons it is necessary to use relatively thin targets in order to prevent deterioration of energy resolution due to slowing of the electrons in the target, except in some cases of particularly high-energy electrons. The most fundamental corrections necessary in order to calculate the differential cross section for excitation from the observed radiation intensities are as follows :

(i) When the number of decay γ-rays is observed, we must add the number of excited nuclei which decay by internal conversion if this is appreciable, and conversely when electrons are observed. This involves the use of theoretical conversion coefficients and knowledge of the decay multipole character.

(ii) When more than one state is excited, the complicated radiation spectrum must be analysed to find how much of the population of each level is due to cascade from higher levels and how much to direct excitation (cf. table 1). Coincidence measurements can help in this analysis.

(iii) When thick targets are used, the observed intensity is determined by an integral of the excitation cross section over the diminishing energy of the projectile along its track. This can be evaluated using the theoretical energy dependence of the Coulomb cross section.

Measurements of differential cross section have been made on the inelastically scattered projectiles by Elbek and Bockelman (1957), using 6 MeV protons on gold. The difficulties of this method are that to maintain energy resolution thin targets must be used, with consequent loss of intensity, and that even small contamination

in the target gives rise to elastic peaks which may be large compared with the inelastic groups (cf. (4.21)). However, the virtue of the method is that it yields the cross section for a given excitation directly, in contrast to the other methods which entail the corrections (i) and (ii) above. The cross section can be calibrated against the elastic scattering, which is given by the Rutherford formula.

7.2. *Optimum Experimental Conditions*

Once we have prescribed the transitions under observation and the experimental method employed, there remain two variable factors—the beam energy T, and the identity of the projectile, which we shall consider in turn.

Beam energy for a specified projectile.

The permissible energy is confined between clear upper and lower limits: at the upper end it should not exceed the Coulomb barrier E_B, in order to keep down nuclear reactions, and at the lower end it must be large enough to prevent the adiabatic parameter ξ (given approximately by (4.12) and exactly by (5.5)) from becoming large; also $T/\Delta E$ should be kept as high as possible. The upper limit E_B is given by (4.17). The lower limit may be measured by the value of T which makes ξ unity, and so from (4.12),

$$T_{\min}(\Delta E) = \tfrac{1}{2}[Z_1 Z_2 m_0^{1/2} e^2 \Delta E/\hbar]^{2/3}$$
$$= 0{\cdot}2 Z_1 (A_1/Z_1)^{1/3} (Z_2 \Delta E_{\text{MeV}})^{2/3} \text{ MeV}. \qquad \ldots\ldots(7.1)$$

$T_{\min}$ increases with ΔE, and so conversely for a given beam energy T there is a highest transition energy $\Delta E_{\max}$ which can be effectively excited, determined from (7.1) by putting $T_{\min}(\Delta E_{\max}) = T$. $\Delta E_{\max}$ is proportional to $T^{3/2}/Z_2$, and for 4 MeV protons on $^{127}_{53}$I, for example, $\Delta E_{\max} = 1{\cdot}9$ MeV.

The working range of beam energies will disappear altogether unless ΔE is small enough that $T_{\min} < E_B$; so from (4.17) and (7.1) we must have

$$\Delta E < 13 \left(\frac{Z_1 Z_2}{A_1 A_2}\right)^{1/3} \text{MeV}, \qquad \ldots\ldots(7.2)$$

i.e. there is an ultimate limit of about 5 MeV to the energies which can be excited. It is further desirable that the experimental conditions validate the semi-classical theory approximately, i.e. we require $\eta > 1$, which will also ensure small Coulomb barrier penetration when $T < E_B$ (cf. (5.15)). We find from (3.8) and (4.17) that $\eta > 1$ whenever the energy T is below the Coulomb barrier, provided $Z_1 Z_2 m_0 > 2\hbar^2/R_0 e^2$, which means numerically $Z_1 A_1 Z_2 A_2^{1/3} > 40$. Hence the semi-classical theory is approximately valid when using protons provided $A_2 > 25$, and when using alphas provided $A_2 > 5$.

When the energy is within the working range, the cross section rises with energy, the more rapidly the larger λ (§ 4.2), and so the intensity is improved by increasing the energy up towards the Coulomb barrier. However, optimum conditions may be determined not by highest intensity so much as by highest signal-to-noise ratio (Alder, Bohr *et al.* 1956, § III). We must consider the following inherent backgrounds, apart from the products of nuclear reactions and the effect of impurities.

(i) When one observes γ-rays, there are, as background, the characteristic x-rays following ionization of atoms by the projectile beam, and there is bremsstrahlung from the projectile in the nuclear field. The former will be important for low energy γ-rays of frequency comparable with that of the x-rays, and also when observing x-rays following internal conversion. The cross section for x-rays following ionization has the approximate theoretical dependence, fairly well confirmed experimentally (Jamnik and Zupančič 1957)

$$\sigma_x \propto Z_1^2(T/A_1)^4 Z_2^{-12}. \qquad \ldots\ldots(7.3)$$

It thus rises more rapidly with energy in the working region than the excitation cross section for any likely multipolarity λ, and so when this background is dominant an energy not much greater than T_{min} is best. In fact, Alder, Bohr *et al.* (1956, § III.B.3) show that for $\lambda = 2$ the ratio signal/x-ray noise is highest when $\xi \sim \frac{1}{2}$, or (from 4.12) $T \sim 1{\cdot}6T_{min}$. The projectile's bremsstrahlung produces a photon background with a continuous spectrum. The theory of this is closely related to that of Coulomb dipole excitation, and the cross section for radiation in the energy range dE_X is given by (Alder, Bohr *et al.* 1956, § II.E.1)

$$d\sigma = (1{\cdot}2.10^{-8})\left(\frac{Z_1}{A_1} - \frac{Z_2}{A_2}\right)^2 Z_1^2 Z_2^2 \frac{A_1 f_{E1}(\eta_1, \xi)}{T_{MeV}} \frac{dE_X}{E_X} \text{ barns.} \qquad \ldots\ldots(7.4)$$

The second factor here is small for all nuclear projectiles except protons, and in fact it is only with protons that this background is found to be appreciable. The ratio signal/bremsstrahlung noise is approximately proportional to $T^{2\lambda-2}$, and so this background does not affect the desirability of high beam energy.

(ii) When one observes conversion electrons, the principal inherent background consists of the continuous spectrum of δ-electrons ejected from atoms by the projectile beam. Alder, Bohr *et al.* (1956, § III.C.2) give the semi-empirical cross section for δ-rays in the energy range dE_δ (measured in kev)

$$d\sigma \sim 2Z_1^2(T_{MeV}/A_1)^4 Z_2^4 E_\delta^{-7}\, dE_\delta \text{ barns.} \qquad \ldots\ldots(7.5)$$

This cross section has the same dependence on T as that for the related emission of background x-rays (7.3), and so the same conclusion follows, viz. the best beam energy is not much in excess of T_{min}. The dependence of (7.5) on Z_2 and E_δ shows that the background is worst for low energy transitions in heavy elements, which is in other respects the best region for internal conversion.

(iii) When one observes the inelastically scattered projectiles, the principal background comes from the elastic scattering, and we see from (4.21) that signal/noise is proportional to $T^{2\lambda-1}$, so high energy is best—also large angles of observation.

Choice of projectile.

Comparing first the ranges of beam energy T needed to produce Coulomb excitation by different projectiles, we see that the upper energy limit E_B is proportional to Z_1, and the lower energy limit T_{min} (7.1) to $Z_1(A_1/Z_1)^{1/3}$. This range of energies is about equally attainable for different projectiles if we have an accelerator of given voltage and completely ionized atoms as projectiles, so that the beam energy is proportional to Z_1. However if only partial ionization of heavier projectiles can be effected, it will more often be impossible to accelerate

them to the desired beam energy than light ones. This can be seen alternatively by considering the maximum transition energy ΔE_{max} which the beam can excite, defined by $\xi = 1$. From (4.12) we see that $\Delta E_{max} \propto (Z_1/A_1)^{1/2}(T/Z_1)^{3/2}$. This is nearly the same for different projectiles provided $T \propto Z_1$ (although the factor $(Z_1/A_1)^{1/2}$ always gives protons a certain advantage), but it will be smaller for heavier projectiles if they can be accelerated less effectively than this. In this connection we may note that He^{2+} ions have been secured by Heydenberg and Temmer (1955), though much work has been done with He^{+}. N^{3+} ions have been used by Alkhazov *et al.* (1956, 1957).

Let us consider the excitation cross section for thin targets. The optimum energy is usually something approaching E_B, and we find from (4.10), (4.11) that if we use different projectiles, making their energies all the same fraction of their respective E_B's, then the cross section is proportional to $Z_1 A_1$, i.e. the heavier projectiles have a much greater cross section.

If thick targets are used, the advantage of heavier projectiles is much reduced, owing to their more rapid stopping. At a very crude estimate, assuming the same beam energy as before, this will reduce the output per projectile by a factor proportional to $(Z_1 A_1)^{1/2}$,† leading to an overall intensity per projectile proportional to $(Z_1 A_1)^{1/2}$, which is still, however, in favour of the heavier particle.

If it is impossible to accelerate a heavier projectile in proportion to Z_1 at the energies in question, the advantage of the heavy projectile in intensity will be greatly diminished and may turn to a disadvantage, particularly with thick targets: specimen figures are given by Heydenburg and Temmer (1956 a, Table III).

A heavier projectile also has an advantage in signal/noise ratio. Penetration of the Coulomb barrier, leading to nuclear reactions, is smaller for heavier particles: if the beam energy is made a fixed fraction of E_B, we find from (3.8) that η is proportional to $(Z_1 A_1)^{1/2}$, and (5.15) shows how this affects the barrier penetration. This somewhat exaggerates the advantage, because we have neglected the effect of the greater size of larger projectiles in lowering the barrier. Of the other backgrounds referred to above, projectile bremsstrahlung is troublesome only for protons. X-rays following atomic ionization and δ-electrons both have a cross section proportional to $Z_1^2(T/A_1)^4$, (7.3) and (7.5), and if we work at an energy which maximizes the signal/noise for these, we readily find that the signal/noise is proportional to $(A_1/Z_1)^{2(\lambda+4)/3}$, which is much to the advantage of particles other than protons. When one observes the inelastically scattered projectiles, it follows from (4.21) that the ratio of the excitation cross section to elastic scattering is proportional to $Z_1 A_1$, provided different projectiles are given energies which are the same fraction of the Coulomb barrier.

We have already remarked that the larger η values of heavy projectiles lead to better validity of the semi-classical theory.

7.3. *Observation of Various Multipoles*

The discussion of § 4.2 has shown that on the single-particle model the cross section would generally be larger the smaller the multipole order of the transition, but in fact dipole excitations are negligible compared with quadrupole, owing to

† This follows from assuming that stopping power $\propto Z_1^2/v$, which results in effective range $\propto (Z_1 A_1)^{-1}$, when $T \propto Z_1$.

the smallness of dipole and the largeness of quadrupole transition moments relative to the single-particle values, and the scarcity of states available for dipole transitions. Although almost all observed transitions are therefore quadrupole, we shall consider the problems involved in observing other multipolarities. The only dipole excitation actually established is that of the first excited state of ^{19}C at 109 kev (e.g. Sherr *et al.* 1954). Octupole excitation in $^{103}_{45}$Rh, $^{107}_{47}$Ag and $^{109}_{47}$Ag has been reported by Jones and Phillips (1957 and unpublished).

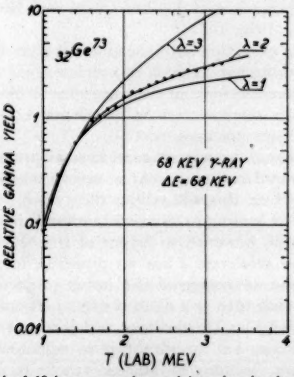


Figure 7. Thin target yield of 68 kev γ-ray observed in α bombardment of $^{73}_{32}$Ge. Dots—experimental results of Temmer and Heydenburg (1954). Full curves—theoretical excitation functions for E1, E2 and E3 transitions, normalized to experimental value at 1·4 Mev. (*After* Alder, Bohr *et al.* 1956.)

First let us consider ways of determining λ from excitation measurements, supposing that it cannot be deduced from known spins and parities.

(i) The excitation function, proportional to $T^{2\lambda-3}f_{E\lambda}(\xi)$ (5.3), is characteristic of λ, but not very sensitively, particularly in the low energy region where the excitation function varies steeply, i.e. when $T \lesssim T_{\min}$ (7.1). This is exhibited in figure 7.

(ii) A method, devised by Bohr and Mottelson, is to excite the transition with two different projectiles at such energies that ξ has the same value for both. Then (4.12) and (5.3) show that the respective cross sections are approximately proportional to $Z_1^2(A_1/Z_1)^{2\lambda/3}$, and so the ratio of the cross sections fixes λ.

(iii) The K/L ratio of the decay conversion electrons following excitation gives information about the multipolarity of the decay and consequently also the multipolarity of the excitation. Likewise the angular distribution and polarization of decay γ-rays, analysed according to § 5.5, involves the excitation and decay multipolarities as well as the spins of the states concerned.

(iv) The angular distribution of the inelastically scattered projectiles depends strongly on λ.

(v) If the radiative lifetime of the excited state is otherwise known we can check whether it is consistent with the $B(E\lambda)$ value derivable for an assumed λ from the Coulomb excitation cross section (cf. (4.16)).

Returning to the case of non-quadrupole excitations of small cross section observed by means of their decay radiation, we note that a difficulty in applying the theory is that the population of the excited state ' A ' may originate almost entirely in the decay of higher states ' B ', ' C ',..., which have been more strongly excited, instead of in direct excitation of the state A. One way of dealing with this is by careful analysis of the radiation spectrum into its several components.

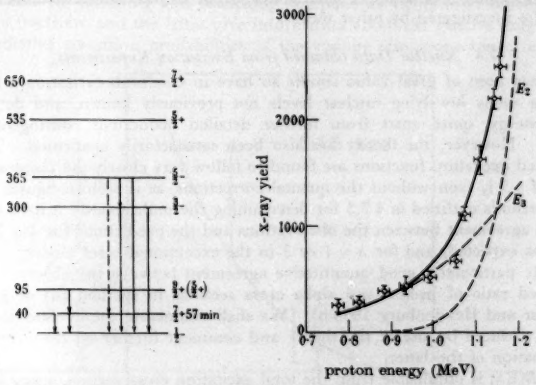


Figure 8. (*a*) Probable level scheme and spin and parity assignments in $^{103}_{45}$Rh. (*b*) Excitation curve for delayed K x-rays produced by bombardment of $^{103}_{45}$Rh with protons. Broken lines are theoretical curves for an E2 excitation of 365 kev, and an E3 excitation of 40 kev. Full curve is sum of broken curves and is approximately fitted to points. The inference is that at low proton energy the 40 kev level is directly excited, and at high proton energy it results mainly from decay of the excited 365 kev level. (*After* Jones and Phillips 1957.)

Alternatively, we can utilize the fact that the steep excitation function for low energy and large ξ, though weakly dependent on λ, is strongly dependent on ΔE (cf. (4.12)). The excitation function for the radiation from state A, if this is excited directly, will be characterized by the corresponding ΔE_A, whereas if it results from decay of a higher state B, the excitation function will be characterized by ΔE_B. This has been used by Jones and Phillips (1957) to establish the existence of octupole excitation in $^{103}_{45}$Rh (figure 8). The importance of decay from higher levels is reduced by lowering the beam energy, since this increases ξ more rapidly for the larger of the ΔE's. Thus in investigating, say, octupole radiation it may be an advantage to use low beam energy, notwithstanding that the cross section rises more rapidly with energy the larger λ, for given ΔE. A further possible complication in applying the theory to a weak excitation is that this may be due in part to second-order effects (§ 5.3), if the final state can be reached through some intermediate state by two strong virtual transitions. This may not be possible to detect by means of the excitation function, in the way that serves for real decay from a higher state.

The detection of a very weak transition may be facilitated by observing the delayed decay radiation, for the excited state tends to be correspondingly long-lived (Jones and Phillips 1957). All background is thus eliminated. This can be particularly helpful with transitions of high multipole order, since, as remarked in § 4.2, the radiative lifetime for a transition increases more rapidly with λ than the Coulomb cross section decreases, *ceteris paribus*. However, by the same token, Coulomb excitation of such states is not so interesting, because they can easily be investigated by other means.

7.4. *Nuclear Data obtained from Excitation Experiments*

It has been of great value simply to have in Coulomb excitation a means of exciting many low-lying nuclear levels not previously known, and determining their energy, quite apart from further detailed deductions contingent on the theory. However, the theory has also been satisfactorily confirmed. Thus the measured excitation functions are found to follow very closely the classical predictions of § 5.1, even without the quantal corrections, as is seen in figures 7 and 8. The methods outlined in § 7.3 for determining the multipolarity λ have exhibited general agreement between the observations and the predictions for $\lambda = 2$ in most cases, as expected, and for $\lambda = 1$ or 3 in the exceptional cases already discussed. We note particularly good quantitative agreement between the observed and the predicted ratio of proton and alpha cross sections in method (ii) of § 7.3 (e.g. Temmer and Heydenburg 1956 b). We shall enumerate the nuclear data which can be deduced by use of the theory and comment further on the experimental confirmation of the latter.

(i) $B(\mathrm{E}\lambda)$ is obtainable from the total excitation cross section, using (5.3). A number of B's have been checked against values deduced from direct lifetime measurements of the short-lived excited states, or from resonance fluorescence measurements, and satisfactory agreement is obtained (e.g. Temmer and Heydenburg 1956 a). The differential cross section for the inelastically scattered projectiles also yields $B(\mathrm{E}\lambda)$, with the aid of (5.6). Elbek and Bockelman (1957 and figure 9) found that the classical theory gives the angular distribution satisfactorily for levels in gold, and their values of $B(\mathrm{E2})$ from this method agreed with determinations from γ-ray measurements.

(ii) Determination of the multipole order λ (§ 7.3) may fix the spin–parity of the excited state.

(iii) From the angular distribution of decay γ-rays the experimental coefficients of the spherical harmonics P_k may be evaluated, and compared with the theoretical values in (5.30), calculated from the Coulomb coefficients $a_k^{\mathrm{E}\lambda}$ and the γ–γ correlation coefficients A_k. The latter depend on the spins and parities of the states concerned, which may consequently be inferred; and if the decay is an E2/M1 mixture, for example, the A_k's are also functions of the mixing ratio δ, which may thereby be evaluated. Possible ambiguities may be removed by likewise studying the γ-ray polarization. When the E2 decay probability $B(\mathrm{E2})$ has been determined from the cross section for excitation as in (i), we can go on to deduce $B(\mathrm{M1})$ from δ. McGowan and Stelson (1957), by investigating pure E2 decays, have obtained experimental values of a_k^{E2} as functions of beam energy, and obtained good agreement with the values given by the quantum calculations. Perturbation

of the angular distribution by extra-nuclear fields is generally negligible, because of the short radiative lifetimes encountered.

(iv) The K/L ratio of decay conversion electrons indicates the decay multipolarity, and, if the decay is an E2/M1 mixture, it determines δ^2. Hence we may be able to obtain B(M1) as in (iii).

(v) When several states are excited simultaneously the analysis of the decay spectrum into its cascading and branching components yields level schemes and branching fractions, and the latter give information about the relative magnitudes of the radiative transition probabilities of the various transitions concerned.

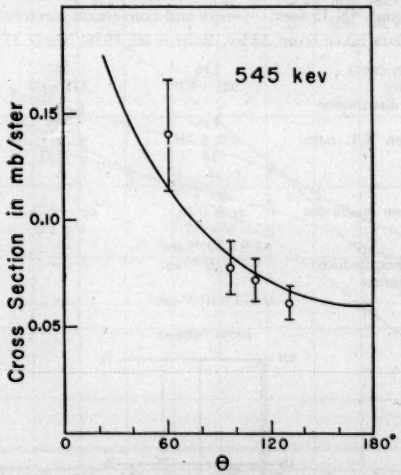


Figure 9. Differential cross section for excitation of 545 kev level in gold by 6 Mev protons. Solid curve is from semi-classical theory, assuming an E2 transition. (*After* Elbek and Bockelman 1957.)

(vi) Comparison of measurements of γ-rays and conversion electrons on the same transition ought to yield experimental determinations of internal conversion coefficients.

Two further points may be remarked about the interpretation of the experiments. Firstly, one may commonly face the problem whether an observed decay radiation originates in direct excitation of the decaying state or follows decay from a higher level as discussed in § 7.3 ; as there explained, we may decide this with the help of the excitation function. Secondly, there is the question whether an observed transition is the result of Coulomb excitation or of nuclear interactions, if the beam energy is high enough to make the latter a likelihood. Detailed investigation of how well the measurements conform to Coulomb theory, for example as regards the excitation function, can decide this. Temmer and

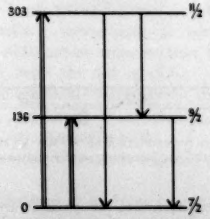
Heydenburg (1956 b) have proved that ordinary Coulomb excitation can still take place at energies high enough for nuclear reactions to occur, by showing that in certain light nuclei the inelastic scattering at energies actually between two resonances is perfectly in accordance with Coulomb theory.

All the information of the above kinds obtained from experiments up to the time of publication is presented in Table IV 2 of the review article of Alder, Bohr *et al.* (1956). As an example, we reproduce in table 1 some data concerning the extensively studied nucleus $^{181}_{73}$Ta.

Table 1. Decay γ-rays from $^{181}_{73}$Ta, excited by protons 1–5 MeV, alphas 3–6 MeV, and ^{14}N 15 MeV. γ-rays and conversion electrons observed.
(Data taken from Alder, Bohr *et al.* 1956, Table IV.2.)

γ energy (KeV)	136	167	303
Decay multipolarity	M1 + E2	M1 + E2	E2
Whether angular distribution measured	Yes	Yes	Yes
Conversion electron K/L ratio	6·5, 6·3†	~7	—
% E2 in decay	16	20	100
Branching ratio	—	—	$\gamma(167):\gamma(303)$ $= 1{\cdot}25-1{\cdot}74$†
Assumed conversion coefficient	$\alpha_L = 0{\cdot}29$, $\alpha = 1{\cdot}7$	$\alpha_K = 0{\cdot}78$, $\alpha = 1{\cdot}0$	$\alpha = 0{\cdot}08$
B (E2) deduced	$1{\cdot}9\times10^{-48}$ cm^4	—	$0{\cdot}57\times10^{-48}$ cm^4
Partial mean lifetime deduced	$0{\cdot}7\times10^{-10}$ sec	—	$0{\cdot}8\times10^{-10}$ sec
B_{sp} theoretical estimate (eqn. (4.15))	$0{\cdot}075\times10^{-48}$ cm^4	—	$0{\cdot}075\times10^{-48}$ cm^4

Level Scheme



† Values obtained from more than one experiment.

§ 8. Excitation by Electrons

8.1. *Experimental Method*

We shall look at the experiments done with electrons of 80–190 MeV from the Stanford linear accelerator (McIntyre *et al.* 1954, Fregeau and Hofstadter 1955, Fregeau 1956, Helm 1956). Inelastic scattering has been observed in association with elastic: the latter has been fully reviewed by Hofstadter (1956), who describes the experimental method. The electrons are scattered off thin targets, then magnetically analysed and detected with a Čerenkov counter, so as to obtain the

angular distribution of the various groups. The absolute cross section is obtained by comparison with scattering from protons, a theoretical cross section being assigned for the latter. However, there are indications that the absolute values may be out by a factor of up to 2. This observation of the scattered projectiles

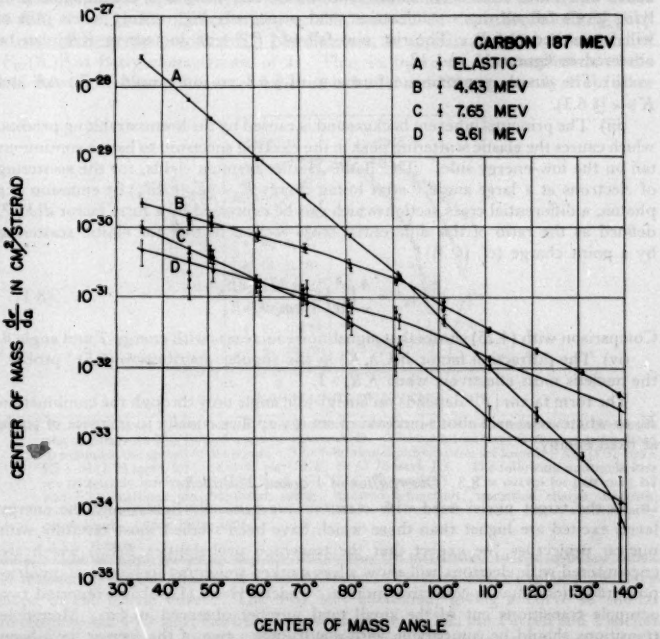


Figure 10. Differential cross sections for elastic and inelastic scattering of 187 MeV electrons by ^{12}C. (*After* Fregeau 1956.)

has the same merits when it comes to interpretation as we have noted in the case of nuclear projectiles, § 7.1. Most work has been done at about 190 MeV, differential cross sections being measured for the excitation of states in the energy range 1–7 MeV in nuclei between ^{9}Be and $^{88}_{38}Sr$, mostly of even A (e.g. figure 10).

8.2. *Optimum Experimental Conditions*

The discussion of § 4.2 shows that in order to get an appreciable cross section when $\lambda \neq 1$ it is necessary to have an energy sufficiently high that KR_0 be not small: we see then from (4.18) that the electron energy should be not much below 100 MeV. The advantages of high electron energy may be listed:

(i) The absolute value of the total cross section (4.13), (4.14) and the ratio $|F|^2$ of the differential cross section for excitation to that for elastic scattering (4.25) both rise with energy at first, the more rapidly the larger λ (except that monopole behaves much like quadrupole § 6.1). However when KR_0 rises much above unity, the cross sections for both elastic scattering and excitation of low-lying levels fall off due to diffraction, and principally high energy levels instead will be excited (§ 6.2). The rise and fall of $|F|^2$ with increasing KR_0 can be observed in figures 11 and 12.

(ii) The simple approximate formulae of § 6.1 are only valid if $T \gg \Delta E$ and $K \gg \kappa$ (§ 6.3).

(iii) The principal inherent background is caused by the bremsstrahlung process, which causes the elastic scattering peak in the electron spectrum to have a continuous tail on the low-energy side. The Bethe–Heitler formula yields, for the scattering of electrons at a large angle θ after losing energy $E_X-(E_X+dE_X)$ by emission of a photon, a differential cross section which can be expressed by a form factor $d|F|_X^2$, defined as the ratio of the differential cross section to that for elastic scattering by a point charge (cf. (6.3)):

$$d|F|_X^2 = \frac{4}{\pi}\left(\frac{e^2}{\hbar c}\right)\ln\left(\frac{2T}{m_e c^2}\right)\frac{dE_X}{E_X}. \qquad \ldots\ldots(8.1)$$

Comparison with (4.25) shows that signal/noise increases with energy T and angle θ.

(iv) The diffraction factor $B(C\lambda, K)$ in the angular distribution (6.5) ' probes ' the nucleus most effectively when $KR_0 > 1$.

The form factor $|F|^2$ depends on energy and angle only through the combination K, so whatever is said about increase of energy applies equally to increase of angle at fixed energy.

8.3. *Observation of Various Multipoles*

As the target nuclei used with electrons are generally lighter and the energy levels excited are higher than those which have been studied most carefully with nuclear projectiles, we expect that the transition probabilities $B(E\lambda)$ which are encountered with electrons will show a less marked weighting in favour of quadrupole transitions due to collective motion. Indeed Helm (1956) has reported two octupole transitions out of the small total number observed so far. Monopole transitions should be comparable with quadrupole: two of the former have been identified. No dipole excitations are known. There are two dynamical factors which render multipoles of high order better observable with electrons than with nuclear projectiles:

(i) The uncertainty in interpreting weak transitions that arises for nuclear projectiles when observing decay radiation, owing to the possibility of its originating from higher levels (§ 7.3), does not exist when observing inelastically scattered projectiles.

(ii) While (4.25) indicates that octupole cross sections should be smaller than quadrupole for smallish KR_0, it happens that for large KR_0 the octupole becomes the greater (see figure 11). The reason is that as KR_0 increases, diffraction cuts down the intensity more rapidly for the smaller λ: this can be recognized in the expansion of j_λ (6.8).

It would be simple to determine the multipolarity λ from measurements, by inspection of the variation of $|F_{E\lambda}(K)|^2$ with K (6.5), if one had reliable observations at values of KR_0 small enough to eliminate diffraction, so that $B(C\lambda, K)$ should effectively take its limiting value for $K \to 0$ (though $\lambda = 0$ and $\lambda = 2$ would not be distinguishable in this way). However, it is most convenient to work at values of KR_0 which make diffraction appreciable, and the K dependence of $|F_{E\lambda}(K)|^2$ is not then exactly predictable. Nevertheless, it appears that the behaviour of $|F_{E\lambda}(K)|^2$ is fairly characteristic of λ. Thus in figure 10, the transitions to the 4·43 MeV and 7·65 MeV levels are believed to be E2 and E0 respectively, from other

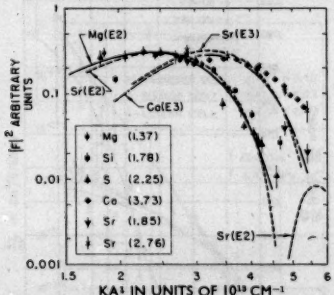


Figure 11. ' Universal curves ' of $|F|^2$ for inelastic scattering of 187 MeV electrons. A composite plot of data from Mg, Si, S, Ca and Sr against $KA^{\frac{1}{3}}$. The measured form factors are normalized to minimize the spread of the points. The following multipolarities are known : Mg (1·37 MeV), E2 ; Si (1·78 MeV), E2 ; Sr (1·85 MeV), E2 ; Sr (2·76 MeV), E3. The following multipolarities are tentatively inferred : S (2·25 MeV), E2 : Ca (3·73 MeV), E3. The curves for particular E2 and E3 transitions are calculated, using ' smeared δ-function ' transition charge densities, with parameters fixed by the elastic cross sections for the respective nuclei. They are arbitrarily normalized. (*After* Helm 1956.)

data. The steeper shape of the E0 curve can be attributed to the quicker onset of diffraction for the smaller λ, as noted above : the same effect explains the steep slope of the elastic curve. The 9·61 MeV excitation, giving a curve so similar in shape to the 4·43 MeV, has been supposed also to be E2, but Ferrell and Visscher (1956) argue in favour of E1. Helm (1956) has remarked that for a number of transitions of known multipolarity in medium-weight nuclei, if we plot $|F|^2$ against $KA^{1/3}$ (equivalent to KR_0), we obtain one ' universal ' curve for $\lambda = 2$ and another for $\lambda = 3$ (figure 11) : using this fact he has inferred further unknown multipolarities.

8.4. *Nuclear Data obtained from Excitation Experiments*

In interpreting experimental results, it is often convenient to deal not with the differential cross section for a process, but the ' measured form factor ' $|F|^2$, defined as the ratio of the measured differential cross section to the calculated differential cross section for elastic scattering by a point charge $Z_0 e$.† We first

† For the latter, more accurate calculations than Born approximation have sometimes been employed.

comment on the experimental verification of the simple Born approximation theory of § 6.1. We have already in §§ 8.2 and 8.3 remarked how the results display the general behaviour expected of them. Born approximation should be obeyed best by light nuclei, but the results up to $^{88}_{38}$Sr exhibit the predicted behaviour well enough. A basic feature of Born approximation is that $|F(K)|^2$ for a given transition depends on T and θ only through K. A plot of Fregeau and Hofstadter's (1955) for a carbon transition (figure 12) shows that this is quite well satisfied.

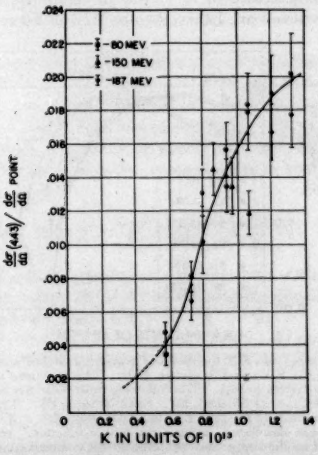


Figure 12. Measured form factor for excitation of the 4·43 MEV level in ^{12}C by 187 MEV electrons, plotted against K. The points obtained at different energies fall nearly on a single curve. (*After* Fregeau and Hofstadter 1955.)

Two obvious rewards of the method are the discovery of new levels and the gleaning of information about spins and parities from the multipolarity λ, provided it turns out that the latter can be reliably deduced (§8.3). One way of marrying theory and experiment is illustrated by calculations on the ^{12}C excitations of figure 10 by Morpurgo (1956), Ferrell and Visscher (1956) and Ravenhall (unpublished), using the shell model. If a definite shell state is assumed for each level concerned, using say, harmonic-oscillator wave functions, then the form factors (6.4) can be calculated completely as functions of K for both elastic and inelastic scattering, and there is only one adjustable parameter, the ' stiffness ' of the oscillator potential, or equivalently the nuclear radius. The latter is adjusted to obtain the best fit with the curve of measured elastic form factor, against angle θ at the actual energy, a further adjustment of scale being allowed because of the uncertainty of the absolute value of the measured cross sections (figure 13). The calculated inelastic form factor can now be compared with the measured (using the same oscillator and scaling parameters) both as function of angle θ and in absolute magnitude.

Good agreement was obtainable, using the shell model, for the angular variation of form factor at 190 Mev for the 4·43 Mev state, but the theoretical absolute value is about half the measured. Ferrell and Visscher (1956) have proposed an admixture of collective motion to raise the cross section. Thus the excitation measurements can be used to test nuclear models.

A more phenomenological analysis has been performed by Helm (1956) on medium weight, even A, nuclei. The elastic form factor can be calculated in the familiar way (Hofstadter 1956) by postulating some form of charge distribution $\rho_1(r)$ (6.16), e.g. a uniform distribution of radius R_0 with smeared-off edge.

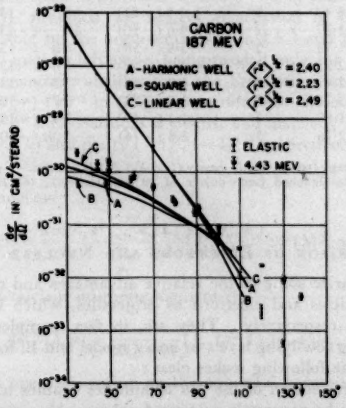


Figure 13. Elastic scattering and excitation of 4·43 Mev level in ^{13}C by 187 Mev electrons. Comparison of experimental and theoretical cross sections, the latter by Ravenhall (unpublished), using shell-model wave functions. Scale factors and r.m.s. radii are chosen to fit elastic data. The excitation is E2. (*After* Fregeau 1956.)

Similarly a transition density $\rho_\lambda(r)$ (6.11) of some shape can be postulated, with adjustable magnitude, and the inelastic form factor calculated. Helm finds that by taking for $\rho_\lambda(r)$ a δ-function, $\delta(r-R_0)$, modified by a smearing factor, and using the same value of R_0 and smearing parameter as for the elastic scattering, he gets a curve of theoretical form factor against angle which agrees well in shape with the experimental result, for some known transitions (figure 11). The absolute magnitude of $\rho_\lambda(r)$ can next be inferred by scaling the calculated form factor to the measured (which is actually itself best calibrated against the calculated elastic). When we have the shape and absolute magnitude of $\rho_\lambda(r)$, then $B(\mathrm{E}\lambda)$ can be calculated from (6.12), or, in the case of monopole transitions, the matrix element of $\Sigma_k r_k^2$ from (6.13). These quantities can otherwise be derived more crudely and simply, using (6.5) and (6.7) or (6.9), by extrapolating the measured $|F_{\mathrm{E}\lambda}(K)|^2$ to small values of K. Helm has evaluated $B(\mathrm{E}\lambda)$ for several E2 and

E3 transitions, in light, even A, nuclei, obtaining results from 1 to 10 times the single-particle values (table 2). They correspond to radiative lifetimes too short to be otherwise readily measurable ($10^{-10}-10^{-13}$ sec). The monopole matrix element (6.13) for the transition to the ^{12}C level at 7·65 Mev has been estimated by Schiff (1955) at about 5×10^{-26} cm^2, which is a little less than the 'single-particle' estimate of § 6.1.

Table 2. Nuclear Data from Excitation by 190 Mev Electrons, after Helm (1956)

(1)	(2)	(3)	(4)	(5)	(6)
$^{12}_{6}$C	4·43	2^+	4·6 (−51)	0·53 (−13)	2·3
$^{24}_{12}$Mg	1·37	2^+	4·6 (−50)	19 (−13)	9·1
$^{28}_{14}$Si	1·78	2^+	3·8 (−50)	6·0 (−13)	6·1
$^{32}_{16}$S	2·25	$(2^+)^a$	4·3 (−50)	1·6 (−13)	5·8
$^{40}_{18}$A	2·4	$(2^+)^b$	1·8 (−50)	3·3 (−13)	1·8
$^{40}_{20}$Ca	3·73	$(3^-)^a$	1·9 (−74)	0·71 (−10)	7·7
$^{88}_{38}$Sr	1·85	2^+	1·3 (−49)	1·4 (−13)	4·6
$^{90}_{38}$Sr	2·76	3^-	8·8 (−74)	1·08 (−10)	7·1

(1) Element; (2) Energy (Mev); (3) I_f, parity ($I_f=\lambda$); (4) $B(E\lambda)$ (cm$^{2\lambda}\times$ power of 10 in brackets); (5) Partial mean lifetime deduced (sec × power of 10 in brackets); (6) $B(E\lambda)/B_{sp}$ (eqn. (4.15)). a, probable, b, assumed.

§ 9. Comparison of Electrons and Nuclear Projectiles

We shall summarize some of the relative advantages and disadvantages of the use of nuclear particles and electrons as projectiles, which we shall denote for brevity Nu and El respectively. They are, in fact, complementary, Nu being best for investigating low-lying levels in heavy nuclei and El for high energy levels in light nuclei, as the following makes clear:

(i) The energy resolution of the Nu techniques enables transition energies of as little as 50 kev to be accurately measured, whereas the energy resolution when detecting 200 Mev electrons is worse than 0·1 Mev.

(ii) The nuclear energy excitable with an Nu beam of given energy T has a severe upper limit ΔE_{max} (§ 7.2), whereas excitation by an electron beam of energy greater than about 100 Mev is subject to no such limit.

(iii) The simple Nu theory (semi-classical) is validated best by large values of η, i.e. large Z_2, whereas the simple El theory (Born approximation) requires small η, i.e. small Z_2.

As regards deduction of the multipolarity λ from observations, the methods available for Nu are varied and reliable (§ 7.3), whereas the El method is as yet uncertain (§ 8.3). On the other hand, El is intrinsically more favourable than Nu for the observation of multipoles higher than quadrupole; only El can deal with monopole transitions; and El should offer the better possibility of observing magnetic transitions (see Appendix).

The accuracy of determination of transition probabilities $B(E\lambda)$ should be quite good with Nu, except when the fundamental corrections of § 7.1 are unduly difficult. The derivation of $B(E\lambda)$ or the monopole matrix elements for El suffers from the uncertainty of absolute calibrations and the indirect analysis involved (§ 8.4).

The El diffraction effects offer in principle more penetrating knowledge about the nuclear wave functions than anything possible with Nu.

Various Nu subsidiary effects, such as the angular distribution of decay γ-rays and K/L ratios of conversion electrons, yield further information as by-products, e.g. M1 transition probabilities.

APPENDIX

Magnetic Excitation

The magnetic interaction is derivable from $V_A^{(1)}$ (2.2) and $V_A^{(2)}$ (2.5) with the vector potential (2.3), provided we neglect retardation of the electromagnetic field, which is justifiable by the arguments used in § 5.3 and § 6.3. The explicit spin interaction $V_A^{(2)}$ may be left aside, being needed only for nuclear projectiles, and being for these generally smaller than the convection interaction $V_A^{(1)}$. On expanding the factor $|\mathbf{r}_p - \mathbf{r}|^{-1}$ in (2.3) as was done for the electrostatic potential (2.9), we can divide $V_A^{(1)}$ into operators of electric and magnetic multipole character, i.e. tensors of order (λ, μ) and parity $(-1)^\lambda$, $(-1)^{\lambda+1}$ respectively, and the magnetic part is:

for r_p outside the nucleus:

$$V_M^{(1)} = \frac{4\pi i Z_1 e^2}{c} \sum_{\substack{\lambda>0 \\ \mu}} \frac{\mathbf{v}_p \cdot \mathbf{L}_p Y_{\lambda\mu}(\theta_p, \phi_p)}{\lambda(2\lambda+1) r_p^{\lambda+1}} \mathfrak{M}^*(M\lambda, \mu) \qquad \ldots\ldots(A1)$$

where $\mathfrak{M}(M\lambda, \mu)$ is a nuclear magnetic moment:†

$$\mathfrak{M}(M\lambda, \mu) = \frac{-i}{ec(\lambda+1)} \int \mathbf{j}_n \cdot \mathbf{L} r^\lambda Y_{\lambda\mu}(\theta, \phi) d\tau, \qquad \ldots\ldots(A2)$$

and $\mathbf{L}$ is an angular operator, acting on the spherical harmonic following it

$$\mathbf{L} = -i\mathbf{r} \times \nabla. \qquad \ldots\ldots(A3)$$

Comparison of (A1) with (2.9) suggests that the essential difference between the formulae for magnetic and electric excitation cross sections of equal order, apart from numerical constants, will be the replacement of the nuclear electric multipole moment by a magnetic one, and the introduction of a factor $(v_p/c)^2$—the latter factor being unity for relativistic electrons. This is indeed borne out by calculation. Thus the semi-classical calculation for nuclear projectiles corresponding to § 5.1, but using the magnetic interaction $V_M^{(1)}$ instead of the electrostatic (2.9), yields the Mλ cross section (cf. 5.3):

$$\sigma_{M\lambda} = \left(\frac{Z_1 e^2}{\hbar c}\right)^2 \frac{B(M\lambda)}{a^{2\lambda-2}} f_{M\lambda}(\xi). \qquad \ldots\ldots(A4)$$

Here $B(M\lambda)$ is defined like $B(E\lambda)$ (4.6), except that the operator in the latter is replaced by $\mathfrak{M}(M\lambda, \mu)$ (A2). $f_{M1}(\xi)$ and $f_{M3}(\xi)$ are plotted in figure 3, and are tabulated by Alder, Bohr *et al.* (1956, Tables II.3 and II.4) and Alder and Winther

† As in the electric case (4.6), our matrix element differs from that of Alder, Bohr *et al.* (1956) by a factor $1/e$.

(1956). As we expect, $f_{M\lambda}$ behaves much like $f_{E\lambda}$, apart from a numerical factor: f_{M1} runs at about $\frac{1}{3}$ of f_{E1} and f_{M2} runs at about $\frac{1}{6}$ of f_{E2}. If we compare the magnetic and electric cross sections of order λ as predicted by the single-particle model (§ 4.2), we see that the magnetic is reduced relative to the electric by three factors:

(i) $(v/c)^2$, which is 0·02 for 10 MeV protons;

(ii) $B(M\lambda)/B(E\lambda) \sim 10(\hbar/Mc\,R_0)^2 \sim 0{\cdot}015$ for $A_2 = 50$;

(iii) a numerical factor from the f's, which is approximately $\frac{1}{3}$ for $\lambda = 1$ and approximately $\frac{1}{6}$ for $\lambda = 2$.

Actual magnetic transition probabilities $B(M\lambda)$ are of the order of single-particle values in favourable cases. An instructive comparison can be made for a mixed M1/E2 transition. If the ratio of M1 to E2 probability for the transition by radiative decay is called δ_{rad}^2, and the ratio of M1 to E2 probability for the inverse Coulomb excitation is δ_C^2, then we see from (4.16), (5.3) and (A4) that

$$\frac{\delta_C^2}{\delta_{rad}^2} = \frac{3}{100}\left(\frac{\Delta E}{m_0 c^2}\right)^2 \eta^2 \frac{f_{M1}(\xi)}{f_{E2}(\xi)}, \qquad \ldots\ldots\text{(A5)}$$

which is evidently extremely small.

The differential cross sections $\sigma_{M\lambda}(\theta)d\Omega$ corresponding to (A4) are expressible by replacing $f_{M\lambda}(\xi)$ by $df_{M\lambda}(\theta, \xi)$: $df_{M1}(\theta, \xi)/d\Omega$ and $df_{M2}(\theta, \xi)/d\Omega$ are plotted in figure 2 and tabulated by Alder and Winther (1956). The quantal calculation for magnetic excitation by nuclear projectiles can be performed by the method of § 5.2, using the interaction (A1) in place of the electrostatic (2.9). The theory is given by Biedenharn *et al.* (1955) and Alder, Bohr *et al.* (1956, § II.B). The result is expressible by replacing $f_{M\lambda}(\xi)$ in (A4) by a function $f_{M\lambda}(\eta_l, \xi)$, and similarly for the differential cross section. Numerical results are given by Biedenharn and Thaler (1956). Ratios of quantal to classical cross sections are shown in figures 4 and 5.

If mixed magnetic and electric excitation can occur, the respective cross sections (both differential and total) are simply additive.

For the angular distribution of decay γ-rays after M1 excitation, the coefficient a_2^{M1} (§ 5.5) is found to be exactly unity, both quantally and classically (Biedenharn and Thaler 1956). After a mixed electric and magnetic excitation, the angular distribution of decay γ-rays contains interference terms arising from the two excitation processes (in contrast to the additivity of the cross sections) which somewhat enhances the small effect of the magnetic excitation: the formulae are given by Biedenharn and Thaler (1956).

The methods in § 7.3 for identifying electrical multipolarity could (singly or in combination) distinguish and identify magnetic multipole excitations. However, it has not been possible so far to observe magnetic excitation with nuclear projectiles (Heydenburg and Temmer 1956 a).

For electron scattering, the effects of the interaction $V_A^{(1)}$ in contributing to electric transitions were considered in § 6.3: we now require its magnetic effects. We follow the same Born approximation procedure as in § 6.1, but using $V_A^{(1)}$ in place of V, and in the nuclear matrix element which corresponds to (6.4) we pick out just the operators which have the tensor character of λ-order magnetic

multipoles (there is no monopole). For $T \gg \Delta E$ and large angles we find (cf. 6.3) for the Mλ cross section :

$$\sigma_{M\lambda}(\theta) = \left(\frac{Z_2 e^2}{\hbar c}\right)^2 \frac{(1+\sin^2\frac{1}{2}\theta)}{8k_1^2 \sin^4\frac{1}{2}\theta} |F_{M\lambda}(I_1 \to I_f, K)|^2, \quad \ldots\ldots(A6)$$

where (cf. 6.5)

$$|F_{M\lambda}(I_1 \to I_f, K)|^2 = \frac{4\pi(1+1/\lambda)K^{2\lambda}}{Z_2^2[(2\lambda+1)\,!\,!]^2} B(M\lambda, K). \quad \ldots\ldots(A7)$$

$B(M\lambda, K)$ is defined like $B(E\lambda)$ (4.6), except that the operator in the latter is replaced by

$$\mathfrak{M}(M\lambda, \mu, K) = \frac{-i(2\lambda+1)\,!\,!}{ec(\lambda+1)K^\lambda} \int \mathbf{j}_n \cdot \mathbf{L} j_\lambda(Kr)\, Y_{\lambda\mu}(\theta, \phi) d\tau. \quad \ldots\ldots(A8)$$

In the limit of small KR_0 we have

$$\mathfrak{M}(M\lambda, \mu, 0) = \mathfrak{M}(M\lambda, \mu), \quad \ldots\ldots(A9)$$

and $B(M\lambda, 0) = B(M\lambda)$ (cf. (6.7)).

The Mλ cross section (A6) is so similar to the Eλ (6.3) as to require little separate discussion. The only significant differences are :

(i) The replacement of $B(C\lambda, K)$ in (6.5) by $B(M\lambda, K)$ in (A7) : if we estimate with the single-particle model, this brings in a factor of the order of $10(\hbar/Mc R_0)^2$ as mentioned above.

(ii) An angular factor $\cos^2\frac{1}{2}\theta$ is replaced by $\frac{1}{2}(1+\sin^2\frac{1}{2}\theta)$.

A method of distinguishing magnetic from electric excitation, devised by Ravenhall, is described by Fregeau (1956), but incorrectly. Two differential cross section measurements of the transition are made at different energies, but at the same value of K, which is effected by adopting two suitably different angles θ. Then according to (6.1), if the transition is electric the two cross sections will be proportional to the respective values of $k^2\cos^2\frac{1}{2}\theta$, and if magnetic to the values of $k^2(1+\sin^2\frac{1}{2}\theta)$.

The exact formulae for magnetic excitation by electrons at lower energies and small angles are given by Schiff (1954) and Alder, Bohr *et al.* (1956, § II.E.3).

The effect of the interaction $V_A^{(2)}$ on M1 excitation by nuclear projectiles has been considered by Biedenharn and Thaler (1956). The cross section for Mλ excitation by $V_A^{(2)}$ is expressible in terms of the electric functions $f_{E(\lambda+1)}(\eta_1, \xi)$ and $df_{E(\lambda+1)}(\theta, \eta_1, \xi)$. The cross sections due to $V_A^{(1)}$ and $V_A^{(2)}$ are additive.

Acknowledgment

I am very much indebted to Dr. A. Bohr and to Professor S. Devons for putting their review articles at my disposal in advance of publication.

References

Alder, K., Bohr, A., Huus, T., Mottelson, B., and Winther, A., 1956, *Rev. Mod. Phys.*, **28**, 432.

Alder, K., and Winther, A., 1953, *Phys. Rev.*, **91**, 1578 ; 1954 a, *Ibid.*, **96**, 237 ; 1954 b, CERN report, T/KA-AW-1 ; 1955 a, *K. Danske Vidensk. Selsk., Mat-fys. Medd.*, **29**, 18 ; 1955 b, *Ibid.*, **29**, 19 ; 1955 c, CERN report, T/KA-AW-4 ; 1956, *K. Danske Vidensk. Selsk., Mat-fys. Medd.*, **31**, 1.

ALKHAZOV, D. G., ANDREYEV, D. S., GREENBERG, A. P., and LEMBERG, I. Kh., 1956, *Nuclear Phys.*, **2**, 65 ; 1957, *J. Exp. Theor. Phys.*, **3**, 964.
BENEDICT, F. D., 1956, *Phys. Rev.*, **101**, 178.
BENEDICT, F. D., DAITCH, P. B., and BREIT, G., 1956, *Phys. Rev.*, **101**, 171.
BENEDICT, F. D., and TICE, G., 1955, *Phys. Rev.*, **100**, 1545.
BETHE, H. A., 1937, *Rev. Mod. Phys.*, **9**, 69.
BIEDENHARN, L. C., and CLASS, C. M., 1955 a, *Phys. Rev.*, **98**, 691 ; 1955 b, *Ibid.*, **100**, 1790.
BIEDENHARN, L. C., GOLDSTEIN, M., MCHALE, J. L., and THALER, R. M., 1956, *Phys. Rev.*, **101**, 662.
BIEDENHARN, L. C., MCHALE, J. L., and THALER, R. M., 1955, *Phys. Rev.*, **100**, 376.
BIEDENHARN, L. C., and ROSE, M. E., 1953, *Rev. Mod. Phys.*, **25**, 729.
BIEDENHARN, L. C., and THALER, R. M., 1956, *Phys. Rev.*, **104**, 1643.
BLATT, J. M., and WEISSKOPF, V. F., 1952, *Theoretical Nuclear Physics* (London : Chapman and Hall).
BOHR, N., 1948, *K. Danske Vidensk. Selsk., Mat-fys. Medd.*, **18**, 8.
BREIT, G., and DAITCH, P. B., 1954, *Phys. Rev.*, **96**, 1477 ; 1955, *Proc. Nat. Acad. Sci., Wash.*, **41**, 653.
BREIT, G., EBEL, M. E., and BENEDICT, F. D., 1955, *Phys. Rev.*, **100**, 429.
BREIT, G., EBEL, M. E., and RUSSELL, J. E., 1956, *Phys. Rev.*, **101**, 1504.
BREIT, G., GLUCKSTERN, R. L., and RUSSELL, J. E., 1956, *Phys. Rev.*, **103**, 727 ; 1957, *Ibid.*, **105**, 1121.
BREIT, G., and LAZARUS, J. P., 1955, *Phys. Rev.*, **100**, 942.
DAITCH, P. B., LAZARUS, J. P., HULL, M. H., JR., BENEDICT, F. D., and BREIT, G., 1954, *Phys. Rev.*, **96**, 1449.
DEVONS, S., and GOLDFARB, L. J. B., 1957, *Handbuch der Physik*, **42**, Ed. Flügge (Berlin : Springer).
ELBEK, B., and BOCKELMAN, C. K., 1957, *Phys. Rev.*, **105**, 657.
FERENTZ, M., and ROSENZWEIG, N., 1954, *Argonne Nat. Lab. Report*, A.N.L. 5324.
FERRELL, R. A., and VISSCHER, W. M., 1956, *Phys. Rev.*, **104**, 475.
FOWLER, G. N., 1956, *Nuclear Phys.*, **1**, 119.
FREGEAU, J. H., 1956, *Phys. Rev.*, **104**, 225.
FREGEAU, J. H., and HOFSTADTER, R., 1955, *Phys. Rev.*, **99**, 1503.
GATTO, R., 1953, *Nuovo Cim.*, **10**, 1559 ; 1955, *Ibid.*, **2**, 669.
GAUNT, J. A., 1927, *Proc. Camb. Phil. Soc.*, **23**, 732.
GLUCKSTERN, R. L., LAZARUS, J. P., and BREIT, G., 1956, *Phys. Rev.*, **101**, 175.
GOLDHABER, M., and WENESER, J., 1955, *Ann. Rev. Nucl. Sci.*, **5**, 1.
GOLDSTEIN, M., MCHALE, J. L., THALER, R. M., and BIEDENHARN, L. C., 1955, *Phys. Rev.*, **100**, 436.
HEITLER, W., 1954, *The Quantum Theory of Radiation*, 3rd edn. (Oxford : Clarendon Press).
HELM, R. H., 1956, *Phys. Rev.*, **104**, 1466.
HEYDENBURG, N. P., and TEMMER, G. M., 1955, *Phys. Rev.*, **100**, 150 ; 1956 a, *Ann. Rev. Nucl. Sci.*, **6**, 77 ; 1956 b, *Phys. Rev.*, **104**, 981.
HOFSTADTER, R., 1956, *Rev. Mod. Phys.*, **28**, 214.
HUBY, R., and NEWNS, H. C., 1951, *Proc. Phys. Soc.* A, **64**, 619.
HUUS, T., BJERREGAARD, J. H., and ELBEK, B., 1956, *K. Danske Vidensk. Selsk., Mat.-fys. Medd.*, **30**, 17.
JAMNIK, D., and ZUPANČIČ, Č., 1957, *K. Danske Vidensk. Selsk., Mat.-fys. Medd.*, **31**, 2.
JEAN, M., and PRENTKI, J., 1954, *C.R. Acad. Sci., Paris*, **238**, 2290.
JONES, G. A., and PHILLIPS, W. R., 1957, *Proc. Roy. Soc.* A, **239**, 487.
LAZARUS, J. P., and SACK, S., 1955, *Phys. Rev.*, **100**, 370.
LEWIS, R. R., jun., 1956, *Phys. Rev.*, **102**, 544.
MCGOWAN, F. K., and STELSON, J. S., 1957, *Phys. Rev.*, **106**, 522.
MCINTYRE, J. A., HAHN, B., and HOFSTADTER, R., 1954, *Phys. Rev.*, **94**, 1084.
MORPURGO, G., 1956, *Nuovo Cim.*, **3**, 430.
MOTT, N. F., and MASSEY, H. S. W., 1949, *The Theory of Atomic Collisions*, 2nd edn. (Oxford : Clarendon Press).

MULLIN, C. J., and GUTH, E., 1951, *Phys. Rev.*, **82**, 141.
SCHIFF, L. I., 1954, *Phys. Rev.*, **96**, 765 ; 1955, *Ibid.*, **98**, 1281 ; 1957, *Nuovo Cim.*, **5**, 1223.
SCHWINGER, J., 1949 a, *Phys. Rev.*, **75**, 898 ; 1949 b, *Ibid.*, **76**, 790.
SHERR, R. C., LI, C. W., and CHRISTY, R. F., 1954, *Phys. Rev.*, **96**, 1258.
SMITH, J. H., 1954, *Phys. Rev.*, **95**, 271.
STELSON, P. H., and MCGOWAN, F. K., 1956, *Bull. Amer. Phys. Soc.*, **1**, 164.
SUURA, H., 1955, *Phys. Rev.*, **99**, 1020.
TEMMER, G. M., and HEYDENBURG, N. P., 1954, *Phys. Rev.*, **96**, 426 ; 1956 a, *Ibid.*, **104**, 967 ; 1956 b, *Ibid.*, **104**, 989.
TER-MARTIROSYAN, K. A., 1952, *Zh. Eksper. Teor. Fiz.*, **22**, 284.
THALER, R. M., GOLDSTEIN, M., MCHALE, J. L., and BIEDENHARN, L. C., 1956, *Phys. Rev.*, **102**, 1567.
WEISSKOPF, V. F., 1938, *Phys. Rev.*, **53**, 1018.

PHYSICS OF NERVE PROCESSES

By P. FATT

Biophysics Department, University College London

CONTENTS

	PAGE
§ 1. Introduction	112
§ 2. Squid giant axon	115
§ 3. Myelinated nerve fibre	132
§ 4. Other material	138
References	142

Abstract. Since 1938 investigations have been made on single nerve fibres and similar material to determine the electrical properties of their surface membrane, which is the structure primarily concerned in excitatory phenomena including the propagation of the nerve impulse. In the resting state the fibre has the characteristics of a leaky capacitative cable. Across the surface of large invertebrate nerve fibres there is a capacity of about 1 μF cm^{-2} in parallel with a resistance of 1000 Ω cm^2 or more. An essential feature of the nerve impulse is the action potential, a transient alteration of membrane potential, which has an amplitude of about 100 mv, lasts about 1 msec at a given position along the fibre, and propagates at a rate of about 20 m sec^{-1}. The membrane resistance decreases during the action potential, while the capacity is unaltered. After the capacitative current has been deducted, the current that flows across the membrane when the potential is displaced from its resting level to a level corresponding to a point on the action potential can be analysed into a component due to a movement of sodium ions and another due to a movement of potassium. The driving force operating on these ions is determined by their concentrations on the two sides of the membrane and the potential difference across it, while the permeability of the membrane toward them varies as a function of membrane potential and time. The rising phase of the action potential is produced by an increase in membrane permeability toward sodium ions, and the falling phase by a delayed increase in permeability toward potassium. In vertebrate nerve fibres that conduct impulses at high speed the process for the generation of the action potential occurs in restricted regions separated by lengths of fibre which serve as cable conductors and which have a thick layer of insulating material at their surface. Action potentials similar to those in the nerve fibre can be generated in some muscle fibres, though not in others. At the junction between a nerve and muscle fibre or between two nerve cells excitatory influences are transmitted by different mechanisms from those involved in the propagation of impulses within the fibre or cell.

§ 1. Introduction

The structural unit for the rapid transmission of messages between distantly separated parts of an animal's body is the nerve fibre, or axon. Messages are conveyed in the form of nerve impulses, signals which travel the length of the fibre at high speed and which can be repeated at short intervals. In the following description of processes in the nerve fibre, attention will be concentrated on electrical events. Apart from the fact that the detection of the nerve impulse has relied to a large extent on electrical techniques, electrical events are of special importance in connection with the mechanism by which the impulse is propagated. The principal observations are of potential differences between different parts of the axon, both at rest and during the propagation of an impulse. The manner of origin of these potentials has in its general features been explained

by R. du Bois-Raymond and J. Bernstein. The interior of the nerve fibre is an aqueous solution similar to that surrounding it, but with a difference in ionic content, and the seat of the potentials is attributed to a membrane separating the two media. In general, as Nernst has indicated, a potential difference will arise where two solutions of different ionic composition come into contact because of differences in the ability of various ions to move in the region of the boundary. In the modern view, by which consideration is given to the effects of individual species of ions, the potential of a solution of composition I with respect to one of composition II will be given by integrating between these limits according to the equation

$$E = \frac{RT}{F} \int_{\mathrm{I}}^{\mathrm{II}} \sum_{n} \frac{t_1 \, dC_1}{z_1 C_1},$$

where the summation is over the n species of ions present in a mixture within the region, z_1 being the valency, C_1 the concentration, and t_1 the fraction of the total conductance (or transport number) for the individual species, and R, T and F are the gas constant, absolute temperature and Faraday's constant. This equation is derived for the condition that there is no current across the boundary, though there will in general be a net movement of ions across it. (It is usually considered that a proper statement of this equation requires the substitution of activities for concentrations, to take account of interactions which modify the ability of ions to diffuse and to add to the conductance. It is doubtful, however, that any use can be made of this since the activities of ion species can be determined in no other way than by potential measurements, and then only as relative quantities; in any case, the potential will be affected only in so far as the ratio of activity to concentration for an individual species varies in the region.)

For the nerve fibre at rest, i.e. not occupied by an impulse, the surface membrane exercises a high degree of selectivity in its resistance to the movement of different ion species with a consequently important effect on the potential difference appearing across it. In a simplified scheme the resting potential that occurs in this circumstance may be attributed to the inability of certain ions to move across the membrane, notably sodium ions which occur in a high concentration in the external medium and in a much lower one inside. Other ions, particularly potassium, are able to move across the membrane and are so distributed that there is no net movement at the resting potential, an equilibrium obtaining in each species of such ions. Owing to the distribution of the non-penetrating ions, there is a potential difference making the outside of the fibre positive with respect to the inside, and the concentration of positively charged penetrating ions such as potassium is higher inside the fibre than outside. The relation which would be expected to hold between the membrane potential E (taken as the potential of the outside with respect to the inside) and the internal and external concentrations (or activities) C_i and C_o for any particular species of penetrating ion, having the valency z, is given by

$$E = \frac{RT}{zF} \ln \frac{C_i}{C_o}.$$

It may be expected that this would be only an approximate description of the resting nerve fibre, since there will be a finite, though small, permeability of its surface membrane toward such ions as sodium. This, combined with an electrical

driving force, measured as the maximum work obtainable from the passage of a quantity of the ion species bearing unit charge, results in there being a continual net movement of these ions across the membrane. An additional movement will occur during the conduction of an impulse. In the animal the essential condition is that there should be no net movement of each such species over long periods of time, so that the composition of the interior of the fibre remains constant. The leakage of ions described above must therefore be balanced by an opposite flow of similar ions, utilizing free energy made available in chemical reaction to overcome the difference in their energy level on the opposite sides of the membrane. The resting potential will thus depend ultimately on chemical processes, though it will be maintained in their absence for considerable periods with little loss provided the membrane has a very low permeability to certain ions.

Characteristic of the nerve impulse is the action potential, during which the potential of the outside with respect to the inside becomes transiently less positive. The free energy made available in the action potential is considered to be derived not immediately from chemical reactions, but from the potential energy implicit in the distribution of those ion species that are not in equilibrium across the membrane in the resting state. The action potential would be generated by an increase in the permeability of the membrane towards these ions resulting in a large movement of ions.

An action potential can be initiated by applying a current pulse between two electrodes in contact with the fibre. The action potential arises during the pulse only at the cathode where current traverses the membrane in the outward direction reducing the membrane potential from its resting level. With a pulse of a given time course, the action potential appears when a critical intensity of current is exceeded. Having been initiated in one region of the fibre, the action potential sweeps along the fibre, as an essential feature of the nerve impulse, with a constant amplitude and velocity. An explanation of the mechanism of propagation of the action potential and thus of the nerve impulse has been given by L. Hermann, and is based on the cable structure of the nerve fibre. The outer conducting medium is presumed to be separated from an inner conductor by a layer of low conductivity which is the site of generation of the action potential. When a segment of fibre is active, current will flow between this and any nearby inactive segment because of the difference in membrane potential between them. Current will flow from the active to the inactive segment along the interior of the fibre and in the reverse direction along the exterior; it will flow inward across the active membrane and outward across the nearby inactive membrane. In this last situation the effect would be equivalent to that of an externally applied current which is capable of exciting. Assuming that in the internal and external media the current is directed everywhere longitudinally, the current flowing across the membrane separating the two media will be given by

$$i_m = \frac{\partial^2 E}{\partial s^2}(r_i + r_o)$$

where E, the membrane potential, may vary with respect to both distance along the fibre s and time t, and r_i and r_o are the resistances of the internal and external media for unit longitudinal distance. Since the action potential propagates at

a constant velocity v, this may also be expressed

$$i_m = \frac{1}{v^2}\frac{\partial^2 E}{\partial t^2}(r_1 + r_0).$$

Both the duration and intensity of current crossing the surface of the fibre are involved in determining the adequacy of a pulse in initiating an impulse. For short pulses a constant charge is required ; as the pulse is lengthened relatively more charge is required, until for long pulses a minimum current is required. There is also a marked spatial and temporal spread in the disturbance produced by a brief pulse, as indicated by its effectiveness in modifying the requirements for a second pulse to excite. These phenomena are presumed to depend on the presence of a large electric capacity and resistance at the surface of the fibre causing a distortion of the applied current pulse.

The foregoing is the classical description of the electrical properties of the fibre, the essential features of which are included in a review by Cremer (1909), where references are given to the early work in the field. Its quantitative development was limited by the experimental use of nerves consisting of bundles of large numbers of nerve fibres. In order to derive the properties of the individual fibres, account would have to be taken of the shunting effect of external fluid as well as of the variation of properties within the population of fibres composing the nerve. The possibility of obtaining rewarding quantitative data did not occur until 1938, when single nerve fibres came to be employed in conjunction with refined electrical techniques. This review is an account of the methods and results of experiments on some types of nerve fibres and other materials of related interest, in which observations have been made directly on the single fibre or cell.

§ 2. Squid Giant Axon

By far the most fruitful material for the elucidation of nerve processes has been the giant nerve fibre from the squid. The squid is a marine invertebrate on the under side of which is a muscular sac which opens to the exterior by a narrow tube. On contraction of the muscle, water is forced out of the tube in a jet, causing a sudden movement of the animal through the surrounding water. The nerve impulses that bring the muscle into action are conducted along part of their course from the brain in fibres of specially large diameter—about 0·5 mm. A uniform length of 5 to 10 cm of giant axon can be dissected. Fine branches that leave it at intervals can be cut without seriously affecting its behaviour.

The properties of the axon are normally studied after its immersion in sea water, the main ionic constituents of which are similar to those of the animal's blood. Analyses have been made of the internal medium of the fibre, the axoplasm, revealing large differences in its composition from that of the external fluid (Bear and Schmitt 1939, Webb and Young 1940, Steinbach 1941, Steinbach and Spiegelman 1943, Keynes and Lewis 1951). The axoplasm of freshly dissected fibres contains sodium, 0·05 mole kg^{-1} ; potassium, 0·32 mole kg^{-1} ; chloride, 0·04 mole kg^{-1}. Sea water contains these ions in the concentrations 0·51, 0·01, and 0·52 mole kg^{-1} respectively. The large excess of these inorganic cations over anions in the interior of the fibre is balanced by organic anions. In general, the distribution of the above ions agrees with the classical view, potassium (and

probably also chloride) being in equilibrium across the membrane for a membrane potential which makes the outside positive with respect to the inside. Sodium ions are evidently not in equilibrium. Furthermore, the nerve fibre gradually loses its internal potassium and gains sodium ; during three hours following isolation this exchange of ions between the axoplasm and the surrounding sea water amounts to about 0·05 mole kg^{-1}.

A study of the electrical characteristics of the squid axon has been carried out by K. S. Cole and his co-workers. In order to determine the impedance of the membrane and of the internal medium a steady current, alternating or direct, is applied between two electrodes in contact with the fluid surrounding the fibre. For alternating current and for direct current pulses of a duration under about 0·1 sec, the electrodes are constructed of a non-reactive metal, usually platinum, and present a large surface in contact with the solution of electrolyte. They operate through the polarization capacity at the metal–electrolyte junction and are effective as long as the impedance of this junction is low compared with that of the remainder of the circuit. The electrodes may be placed either transversely or longitudinally to the fibre. In the first case the nerve is placed in a channel of square cross section, two opposite sides of which are occupied by electrodes. Since the parts of the isolated axon near the cut ends will be in an injured, and therefore altered, condition, it is necessary for the axon to extend in both directions beyond the electrodes. A complication will thus arise at the edges of the electrodes where there will be a longitudinal component of current, the effect of which can be reduced by having a relatively long stretch of nerve between the electrodes. Impedance is measured with an alternating current Wheatstone bridge (Curtis and Cole 1938).

With the measuring cell large enough to allow the axon to be conveniently placed within it, it is found that at low frequencies nearly all the current flows around the fibre through the external fluid, while the appreciable fraction of the applied current that does enter the fibre at high frequencies is carried almost entirely by the capacity of the membrane. Measurements at high frequencies will allow the internal conductivity to be calculated, but this quantity can be determined more accurately by the use of longitudinal electrodes, there being no indication that the interior of the fibre is not isotropic in its conductivity. Although the conductance of the membrane cannot be measured with transverse electrodes, the transition in the observed impedance between high and low frequencies allows precise information to be obtained about the membrane capacity. The geometry of the measuring cell is taken into consideration by the use of an equation of Rayleigh

$$\frac{1-R_o/Z}{1+R_o/Z} = \rho\frac{1-R_o/Z_e}{1+R_o/Z_e}$$

which relates the impedance for unit length of cell Z to the specific impedance of the fibre Z_e (occupying the fraction of the volume of the cell ρ) and that of the material filling the remainder of the cell R_o (in this case a pure resistance). This equation, which is similar to one derived by Maxwell for a suspension of parallel cylinders, applies to the situation in which there is a transverse flow of current across a cell of square cross section containing a circular cylinder with a diameter appreciably less than the side of the square. The impedance offered by the nerve

fibre will be contributed to by the resistivity of the internal medium R_i and the specific impedance of the surface membrane Z_m. This may be taken into consideration by replacing Z_c in the above equation by $(R_i + Z_m)/a$ where a is the radius of the fibre.

For convenience of interpretation the Wheatstone bridge measurement at each frequency is expressed as a resistance and reactance in series $Z = R_s + iX_s$. From a number of measurements at various frequencies an impedance locus is drawn having the form shown in figure 1. The range of frequencies over which the variation in Z occurs depends on the degree of shunting of the fibre by the external fluid ; in a cell of convenient size this extends from about 10^3 to 5×10^5 c/s.

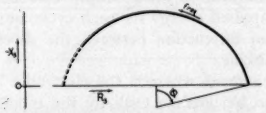


Figure 1. Form of impedance locus obtained for electrodes placed on opposite sides of nerve fibre. Membrane phase angle is given by ϕ.

It may be shown that for a network consisting of a capacity with resistances in series and in parallel the impedance locus will be the arc of a circle whose centre lies on the axis of R_s. The experimental data approach fairly closely to this relation, though deviating from it in a consistent way. A better fit is obtained with the arc of a circle with its centre lying in the region of positive X_s. A relation of this type would arise if the frequency dependent component in the network, considered provisionally to be a capacity, were an impedance of constant phase angle between 0 and $-90°$. The phase angle will then be given by half the angle subtended at the centre of the arc by the points where the arc meets the axis of R_s. For the squid axon this is $-76°$. Extrapolation to the R_s axis also gives the limiting resistances for zero and infinite frequencies. With this information and the observed complex value of Z at a given intermediate frequency it is possible to work out what the capacity in the network must be, the equation of the geometry of the measuring cell allowing this to be expressed as the capacity for unit area of membrane C_m. This procedure, which ignores the fact that Z_m cannot be represented exactly by a capacity and resistance in parallel, appears to be justified, since only a very small frequency dependent resistance in series with C_m will suffice to give the observed phase angle. The membrane capacity may itself however not be entirely independent of frequency, though over the range of frequencies where this component will be significant in the behaviour of the nerve fibre (considering the resistances in series and in parallel with it in any given situation), its variation is probably negligible. The value of C_m is calculated to be 1·1 μF cm^{-2}.

This capacity yields information on the material composing the membrane. Physical and chemical techniques reveal at the surface of various types of nerve fibres one or more layers composed of fat and protein about 70Å thick. (With the electron microscope several layers are observed at the surface of the squid axon (Geren and Schmitt 1954.) It is probable however that only the innermost one contributes to the measured membrane capacity, the others being shunted by a

relatively low resistance provided by a gap filled with external fluid which leads from the outer surface of the inner layer to the exterior.) The average dielectric constant D for the membrane, calculated from the above values for capacity C_m and thickness d by the relation $C_m = D/4\pi d$, is found to be about 10. It may be inferred from this high value that there exist within the membrane molecules with permanent electric dipoles which are rotated by changes in electric field. The presence of fat and protein molecules with ionized side chains, which is indicated on chemical evidence, would provide the necessary polarizability. A constant phase angle between 0 and $-90°$ as found in the nerve fibre membrane is also commonly observed in other dielectric materials which are not perfectly homogeneous. This behaviour is most simply described as a dissipation of a constant fraction of the applied energy for each cycle, and in the present case this loss probably arises from interaction between the dipolar molecules and other constituents of the membrane.

The alternative arrangement whereby the electrodes are placed longitudinally along the nerve fibre can be used to evaluate the membrane resistance and the internal resistance of the fibre. Instead of the fibre being enclosed in a channel as previously described, it is suspended in paraffin oil or moist air, under which conditions a thin, fairly uniform layer of sea water adheres to it. To analyse the data obtained by this method, Kelvin's theory of the cable is applied on the assumption that the internal and external media can be represented as longitudinal conductors with a leakage of current across the membrane separating them. The solution for the impedance Z between two electrodes s distance apart in contact with the exterior of the fibre which extends effectively for an infinite distance in each direction (i.e. beyond the distance to which current will spread in the longitudinal conductors) is given by

$$Z = \frac{r_i r_o}{r_i + r_o} s + \frac{2 r_o^2 \lambda_z}{(r_i + r_o)(K + \coth s/2\lambda_z)}.$$

In this equation, r_i and r_o are the internal and external resistances for unit length of fibre; λ_z is defined as $[z_m/(r_i + r_o)]^{1/2}$ where z_m is the membrane impedance for unit length of fibre. The term K depends on electrode conditions: it is unity when the electrode makes contact over a very short length of fibre, and $[(r_i + r_o)/r_i]^{1/2}$ when the electrode makes contact over a long length of fibre, abolishing any potential gradients in the external conductor outside of the region between electrodes.

Alternating current measurements made with this arrangement of electrodes are noteworthy in that the impedance of the membrane can be observed for lower frequencies than those used with transverse electrodes, the variation in Z occurring below about 10^4 c/s (Cole and Baker 1941 b). The impedance locus obtained by this method, using a large electrode separation, has the form shown in figure 2. At high frequencies the membrane capacity will cause z_m to be small and the behaviour of the fibre will be described by the first term on the right in the above equation, representing simply the internal and external conductors in parallel, and $r_i r_o/(r_i + r_o)$ may thus be evaluated. For very low frequencies and direct current z_m will be the membrane resistance for unit length of fibre r_m. When s is made large both s and z_m will disappear from the denominator and the observed Z at any frequency, by comparison with the limiting values for zero

and infinite frequencies where it is a pure resistance, can be transformed into $z_m/r_m = r_s/r_m + ix_s/r_m$, which pertains to the fibre membrane. Although these measurements (obtained with s fixed and only the frequency varied) do not allow a separate evaluation of membrane capacity or resistance, their product, the time constant of the membrane, can be obtained from x_s/r_m at a given frequency. It agrees in its order of magnitude with the value calculated from the separate determination of capacity and resistance by other methods. Considered in polar coordinates, the above transformation requires that the phase angle of the membrane impedance z_m should be given by one half the angle of approach at high frequencies

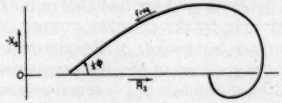


Figure 2. Form of impedance locus obtained for electrodes placed a long distance apart along nerve fibre.

of the locus of Z to the R_s axis as indicated in figure 2. The value of the phase angle obtained in this way from measurements made with frequencies of the order of several thousand cycles per second, is approximately $-75°$, in agreement with the result derived by the use of transverse electrodes. As the frequency is lowered the behaviour of Z is not consistent with z_m, consisting of a parallel arrangement of a resistance and a capacity with dielectric loss. At frequencies below about 200 c/s, X_s frequently appears positive, and variation of the electrode separation at a single frequency indicates that the membrane is responsible. The form taken by the impedance locus can be approximated by an inductance having a value of about $0{\cdot}2\,\text{H cm}^{-2}$ in series with the membrane resistance and in parallel with the capacity. It appears however that this electrical behaviour at low frequencies arises as a result of the membrane resistance varying with the membrane potential after a short delay (Cole 1941, Hodgkin and Huxley 1952 d).

Measurements made with direct current and with varying electrode separation have been used for determining the membrane resistance and internal resistivity of the fibre. The three quantities r_i, r_o and r_m referring to unit length of fibre are readily obtained from the relation between s and the observed resistance. The value of r_o gives an indication of the thickness of the layer of sea water adhering to the outside of the fibre under the conditions of the experiment; this comes out at about 0·05 mm for a nerve suspended in oil. The values of r_i and r_m are converted to the corresponding specific resistances of the internal medium and the membrane by the relations $R_i = r_i/\pi a^2$ and $R_m = r_m/2\pi a$. The value obtained for R_i is 30 Ω cm—about 1·4 times the resistivity of sea water. The higher resistivity of axoplasm than of sea water is attributable to the replacement of chloride by larger organic anions. The value of R_m is computed to be about 1000 Ω cm². Considered with a probable thickness of 70Å this indicates that the material composing the membrane has a resistivity of an order of 10^7 times that of sea water.

In the experiments described above the voltage impressed on the membrane is kept as small as is consistent with accurate observation to avoid disturbing the

mechanism by which the action potential is produced. Furthermore, with impedance measured between two electrodes placed on parts of the fibre which are under similar conditions, a non-linearity in the conductance element may complicate the observations without its real nature being revealed in the analysis. An arrangement allowing the behaviour under the anode, where current is entering the fibre, to be distinguished from that under the cathode, where current is leaving, employs transverse electrodes for impedance measurement which, with suitable decoupling of the two electrodes, are also included as either cathode or anode in a circuit for passing direct current, this latter circuit being completed by a third electrode placed a long distance away along the fibre. The transverse electrodes are made narrow so that the region of fibre included in the impedance measurement is affected relatively uniformly by the constant current. The resultant increased significance of a longitudinal component of the alternating current at the edges of the electrodes reduces the accuracy of the measurements. With the Wheatstone bridge initially balanced, the application of a constant current, making the transverse electrodes either anode or cathode, produces an unbalance. A steady condition is attained several milliseconds from the onset of the current. From the requirements for rebalancing the bridge during a constant current it appears that the impedance change can be fully accounted for by a change in membrane resistance to the small alternating current superimposed on it in the region of the transverse electrodes. The differential, or dynamic, resistance thus determined decreases under the cathode and increases under the anode.

Variations of membrane resistance are more satisfactorily determined from observation of changes produced in the membrane potential on the application of constant current pulses of graded intensity (Cole and Curtis 1941). As in the experiment described above, a constant current is applied between two electrodes spaced far apart along the fibre. For recording membrane potential changes, potentials are led from the interior of the axon by means of a glass capillary with a diameter not more than about $\frac{1}{4}$ that of the axon and filled with a conducting material. This electrode is pushed along the axis of the fibre until its tip, at which electrical contact is made with the surroundings, is in an uninjured part of the axon several millimetres or more from its point of entry. Potential differences are recorded between this and another electrode in the external fluid immediately opposite it, where one of the current electrodes is also situated. When a current step is applied, the potential recorded across the membrane takes several milliseconds to reach a steady level. Under the anode the potential change has approximately the time course to be expected for a cable with the distributed properties already indicated, i.e. it is proportional to $\mathrm{erf}(t/\tau)$, where τ is the time constant of the membrane given by $R_m C_m$. Under the cathode, when the current is sufficient to produce a potential change of a few millivolts, the potential displays an early peak, which for current beyond a critical intensity develops into a propagated action potential. The voltage–current relation for the steady condition occurring beyond a few milliseconds from the start of the current has a marked curvature, the change in potential produced by a given current applied between the two longitudinal electrodes being greater under the anode than under the cathode. From this curve, which depends on the fibre acting as a cable, extending effectively an infinite distance in both directions, the relation that holds at every

point along the fibre between membrane potential E and membrane current density I_m can be derived by means of the equation

$$I_m = \frac{r_o^2 I}{4(r_i + r_o)} \frac{dI}{dE}$$

where I is the current applied to the fibre and dI/dE refers to changes occurring under the electrode. Around the position of the resting potential (at zero I_m) the differential resistance of the membrane (dE/dI_m) changes by a factor of 100 for a change in E of about 20 mv. In either direction outside this range of potential the differential resistance rapidly approaches a limiting value. The possibility has been considered that this rectification might arise from a difference in the concentration of penetrating ions, particularly potassium, on opposite sides of the membrane (Goldman 1943). Although this problem cannot be dealt with exactly without information on the conditions imposed on the movement of ions at each point through the membrane, it appears probable from a simple model which assumes a linear potential gradient that this factor could make only a very small contribution to the observed high degree of rectification. In fact the rectification appears to be attributable to a variation in the permeability of the membrane to potassium ions, under the control of the membrane potential.

The technique of placing an electrode in the interior of the axon has been used to make accurate measurements of the resting and action potentials (Curtis and Cole 1942, Hodgkin and Huxley 1945). In order that the potential difference recorded between the two electrodes inside and outside the axon should be stable over long periods of time, the electrolyte–metal junctions in the leads to the recording apparatus are contrived so as to pass very small currents in a thermodynamically reversible manner through an oxidation–reduction type of reaction. The arrangement most generally used is to make contact with the nerve fibre through a solution of chloride ions in contact with a silver chloride coated silver wire leading to the apparatus. The attempt is also made to eliminate or correct for all asymmetry in the potential differences occurring at the various junctions between the two inputs to the recording apparatus except for that at the surface of the axon. In fact there will be an uncertainty of several millivolts in the potential difference that occurs at the tip of the internal electrode where the solution in the glass capillary makes contact with the solution in the interior of the axon. The resting potential is found to be about 60 mv, the outside positive with respect to the inside. An action potential set up at a distant position along the fibre and propagated into the region of observation is found to have an amplitude of about 110 mv, the potential of the outside relative to the inside changing in the negative direction. During the action potential therefore the membrane potential is not only reduced towards zero, but is reversed by about 50 mv. This is illustrated in figure 3. The amplitudes of the resting and action potentials recorded by this method are about twice those which can be recorded between two distantly separated electrodes on the outside of the fibre, one of which (the reference electrode) is on a region of fibre in which the membrane potential is reduced to zero, usually by treating the exterior of the fibre with potassium chloride solution. The potential difference thus recorded will be reduced by the factor $r_o/(r_i + r_o)$ compared with the membrane potentials occurring in the region of normal axon.

The action potential begins with a phase of increasing slope during which current flows outward across the membrane as the active region approaches the position on the fibre where the membrane potential is observed. The rapid part of the action potential has a duration of about 1 msec at 20°c, of which the rising phase accounts for about 0·25 msec. This is followed by a more prolonged phase of increased membrane potential, with a maximum at about 15 mv above the resting level and a total duration of about 10 msec. From external recording between two points along the fibre, the conduction velocity is determined to be 20 m sec^{-1} (Schmitt and Schmitt 1940). The wavelength for the rapid potential change will therefore be about 20 mm.

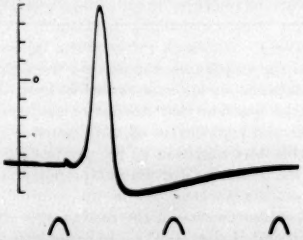


Figure 3. Time course of action potential recorded between an electrode in the interior of squid axon and another in the external solution. Zero potential is obtained with both electrodes in the external solution. Vertical scale is in 10 mv; marks along bottom are 1 msec apart. (*After* Hodgkin and Huxley 1939.)

The impedance of the fibre during the passage of an action potential has been observed, using transverse electrodes and alternating current (Cole and Curtis 1939). The electrodes for the impedance measurement are made narrow compared with the wavelength of the action potential to obtain a satisfactory temporal resolution of alternations in impedance. With the electrodes connected to a Wheatstone bridge which is balanced when the fibre is at rest, the propagation of an action potential past them throws the bridge temporarily out of balance. It is concluded from the requirements for having the bridge balanced at a particular time during the action potential that the impedance change consists entirely in a decrease of membrane resistance, the membrane capacity and phase angle remaining unaltered. At its maximum the change in the resistance of the membrane amounts to a reduction to about 30 $\Omega\ cm^2$. The time course of the bridge unbalance, which is very nearly proportional to the change in conductance, is shown in figure 4, along with the action potential recorded at the same place (with respect to a distant inactive portion of fibre) under similar conditions of timing. The beginning of the impedance change is seen to coincide approximately with the inflection on the rising phase of the action potential, i.e. with the reversal of membrane current from outward to inward. When corrected for the finite conduction time along the length of fibre affected by the measuring current and for the lag in response of the recording apparatus, the time of rise of the impedance change is determined to be about 0·1 msec. The maximum impedance change coincides approximately with the summit of the action potential. The falling

phase is more prolonged than that of the action potential, and extends into the phase of increased membrane potential.

The influence of temperature on the resting and action potentials has been examined (Hodgkin and Katz 1949 b). The resting potential remains practically constant over the range 0 to 20°c ; it diminishes with further increase of temperature up to about 35°c, beyond which there is an irrecoverable loss of resting potential and failure of propagation of the action potential. The peak amplitude of the action potential also remains constant or decreases slightly with increasing

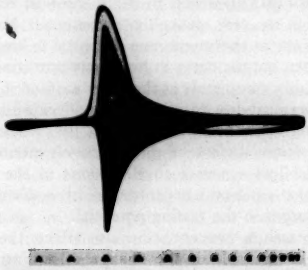


Figure 4. Time course of a.c. bridge unbalance recorded on propagation of an impulse along a squid axon with electrodes placed on opposite sides of it. A record of the action potential, obtained with comparable timing, is superimposed. Marks along bottom are 0·5 msec apart. (*After* Cole and Curtis 1938.)

temperature from 0 to 20°c. Its maximum rates of rise and fall on the other hand vary steeply ; for every 5°c rise in temperature in this range the rate of rise increases by about 40% and the rate of fall increases by about 70%. The very small variation in resting potential and peak amplitude of the action potential over the normal range of temperatures for this material, suggests that the sources of e.m.f. that determine these values do not depend immediately on chemical processes, the rates of which would be expected to vary greatly with temperature. The considerable effect of temperature on the rate of change of potential suggests that the conductance alternations whereby these changes take place do involve chemical processes.

The effects of the ionic environment on the resting and action potentials have also been investigated (Curtis and Cole 1942, Hodgkin and Katz 1949 a). When the potassium concentration is increased above about 40 mmole $l.^{-1}$, which is three times the concentration in sea water, the curve of the membrane potential against the logarithm of the potassium concentration is a straight line with a slope of about 50 mv for a ten-fold change in concentration. This conforms closely with the theoretical relation for a membrane conductance due entirely to potassium ions, i.e.

$$E = \frac{RT}{F} \ln \frac{C_i}{C_o}$$

with the internal concentration C_i being constant. Moreover the position on the curve for zero membrane potential indicates that the internal concentration is

about 400 mmole $l.^{-1}$, which agrees with chemical determinations. With external concentrations less than 20 mmole $l.^{-1}$ the curve is less steep and flattens out to a potential about 5 mv above the resting potential for low concentrations. This is interpreted as the effect of a finite permeability towards other ion species which are present in significant amounts, and which contribute an increasingly greater fraction of the conductance in the outer part of the membrane as the external potassium concentration is reduced. However, the observed curvature is probably greater than if the permeability of the membrane to the various species of ions were constant, and the rapid approach to the theoretical relation with increasing potassium concentration is very probably accentuated by an increase in the permeability to potassium as the membrane potential is lowered. Another effect likely to operate to flatten out the curve at low potassium concentrations is a failure of the external fluid to mix completely at the outer surface of the membrane. The condition of a finite permeability to other ions will cause a net outward flux of potassium at low external concentrations, and the resultant accumulation of potassium ions at the outer surface of the selectively permeable membrane will then make the potential less sensitive to alterations in the bulk of the external fluid. Variations in the external concentration of sodium on the other hand produce only small changes in the resting potential.

Variation in the potassium concentration also affects the height of the action potential, presumably as a result of the change of resting potential. Of more significance to the mechanism of production of the action potential is the dependence of its height on the external sodium concentration. Over the range of concentrations from $\frac{3}{2}$ to $\frac{1}{2}$ that of sea water, a plot of the peak amplitude of the action potential against the logarithm of the sodium concentration gives a straight line with a slope of about 55 mv. Since the resting potential does not vary appreciably, the same relation will apply to the membrane potential at the summit of the action potential. It is the relation that would be expected if the conductance of the membrane at the summit of the action potential were due mainly to sodium ions. Their internal concentration can be calculated by extending the relation for the height of the action potential to the point at which the action potential will just balance the resting potential to give zero membrane potential. The value of about 50 mmole $l.^{-1}$ thus obtained is in agreement with chemical determinations. Although the action potential does not fail until the external sodium concentration is reduced to below about 25 mmole $l.^{-1}$ ($\frac{1}{20}$ of the concentration in sea water) the relation described above does not apply down to such low concentrations, the changes occurring below about 200 mmole $l.^{-1}$ being greater than predicted. This may be accounted for by assuming that the sodium permeability is a function of membrane potential, so that when the action potential is small the contribution to the membrane conductance made by sodium ions is not sufficient for it to determine the behaviour of the potential. In addition to these effects observed within seconds or minutes after changing the composition of the external solution, there are long term ones as a result of net movements of ions into and out of the fibre producing a gradual change in the internal composition.

The extension of these investigations has depended upon elaboration of the technique of the internal electrode making it possible to observe the behaviour of the membrane when acted on uniformly by an applied current (Marmont 1949,

Hodgkin, Huxley and Katz 1952). The internal electrode has been modified, so that the electrical contact is not restricted to the tip of a capillary, but extends for a considerable distance at the exposed surface of a wire passing along the axis of the fibre. Since in this electrode there is no solution of known composition interposed between the internal medium and the electrolyte–metal junction, a satisfactory measurement of the resting potential cannot be obtained. In reducing observations on displacements of potential to absolute levels of membrane potential, it is necessary to employ the resting potential measurement made with an electrode of the capillary type. In most of the experiments described below, two long electrodes are placed side by side in the interior of the axon, current being applied through one and potential recorded with the other. This avoids recording the potential developed across the electrolyte–metal junction when a current is passed. Since only a fraction of the axon can be used in observing the membrane properties and it is desirable that there should be no current between the portion under observation and adjacent parts, a system of guard rings is employed to keep the membrane potential uniform over a length of fibre extending in both directions beyond the region in which the potential recording electrodes are situated. This involves placing insulating partitions in the external medium to isolate electrically the central part of the axon, where potentials are recorded, from the peripheral parts. A uniform current is then applied to central and peripheral parts. In a more complicated system which is effective when the fibre properties are not uniform between the central and peripheral parts, the current applied to the membrane in the periphery is controlled to eliminate any potential difference in the external medium between them.

With the membrane acted on uniformly by the applied current, short pulses in either direction cause proportional changes in the instantaneous value of the potential due to charging of the membrane capacity. From the size of the displacement and the amount of charge applied in the pulse, C_m is found to have a value of 1 μF cm^{-2}, which is in agreement with that obtained from alternating current measurements. Furthermore, following an inward pulse which produces a hyperpolarization (i.e. an increase of membrane potential), the potential returns to its original level exponentially, being proportional to $\exp(-t/\tau)$, where τ is the time constant of the resting membrane. An action potential is elicited when an outward pulse depolarizes the membrane (i.e. reduces the membrane potential) beyond a critical amount of about 12 mv. Close to this, the membrane potential remains fairly steady for a period of up to several milliseconds after the initial displacement before either returning to the resting level or developing into an action potential. Since no current flows between different regions of the fibre and the applied current has ceased to flow before the action potential it sets up has developed appreciably, the response of the membrane occurs in the absence of current. The action potential recorded under these conditions is very similar to the propagated action potential past the initial phase of increasing slope in the latter, from which it is inferred that the impedance of the active membrane is so low that, in the case of propagation, the current flowing across it, because of the neighbouring inactive parts of the fibre, is insufficient to cause an appreciable alteration of membrane potential.

More detailed information on the electrical properties of the membrane involved in the development of the action potential is derived from the manner

in which the potential alters a short time after the application of a brief pulse. In the absence of membrane current, a change in charge on the capacity can be effected only by a net movement of ions across the membrane. The behaviour of the membrane will therefore be described by $I_i - C_m\, dE/dt = 0$, where I_i is this ionic component of the current. Provided that C_m is unaltered, I_i can be determined from observation of the rate of change of potential. The latter has been observed a few tenths of a millisecond after the application of pulses of various amplitudes, this being the time at which a large outward pulse will bring the membrane into the state of activity represented by the summit of the action potential. Values of I_i, when plotted against the membrane potential at which they occur, give a curve of the form indicated by the full line in figure 5. It is

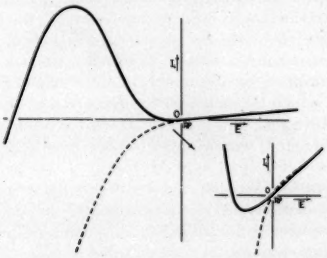


Figure 5. Form of curve relating non-capacitative (or ionic) current to membrane potential a few tenths of a millisecond after a sudden displacement of potential from the resting level indicated by full curve. Relation in steady state is indicated by broken curve. Portions of these curves around the resting potential (RP) are given with altered scales in lower right-hand corner.

seen that both I_i and dI_i/dE vary continuously with E. When the membrane is hyperpolarized (E is made greater than the resting potential), dI_i/dE is nearly constant at about $0{\cdot}0005\ \Omega^{-1}\ cm^{-2}$. Around the level of the resting potential dI_i/dE decreases with decreasing E. As the membrane is depolarized (E is made less than the resting potential), dI_i/dE is reduced to zero, and is negative for depolarizations between 7 and 70 mv. For depolarizations between 12 and 110 mv, the ionic current is inward (I_i is positive), and will tend to depolarize the membrane further, corresponding to positions on the rising phase of the action potential. The critical position for setting up an action potential, at a depolarization of 12 mv, corresponds to the position where I_i is zero and dI_i/dE is negative. The summit of the action potential at a depolarization of 110 mv corresponds to the position where I_i is zero (since dE/dt is zero) and dI_i/dE is positive. The latter is found there to give the high value of differential conductance of about $0{\cdot}03\Omega^{-1}\ cm^{-2}$, which agrees with the finding of alternating current measurements.

In order to examine the relation between I_i and E at any time a device has been employed which allows the voltage to be set at any value, by providing current from a feedback network which includes the potential recording electrodes together with a source of voltage pulses for modulation. When the membrane is subjected to a rectangular voltage pulse, the current in the membrane capacity consists of a

surge at the beginning and at the end of the pulse ; during it, the current across the membrane will be entirely the ionic component I_i. Besides that due to instrumental limitations, a slight lag occurs in the capacitative current, owing to the phase angle of the membrane not being exactly $-90°$. A more disturbing effect is caused by the presence in the internal and external media of small resistances in series with the membrane. In this circumstance, when the applied current varies with time, there is a variation of the membrane potential as a result of the extra resistance which is in series with the membrane capacity. The feedback network can be made to compensate for the effect, though the major results obtained with the technique are not dependent on this.

The curve obtained for the current flowing a few tenths of a millisecond after the start of a voltage pulse is the same as that in figure 5, confirming that the membrane capacity is constant in the determination of I_i from the rate of change of potential. Beyond several milliseconds from the start of the pulse, the relation between I_i and E follows the curve shown by the broken line in figure 5. This is similar to the relation obtained from the analysis of measurements on the fibre with cable properties as discussed earlier. Examination of the time course of membrane current following a sudden displacement of membrane potential from the resting level reveals the manner in which the transition takes place between the two types of behaviour indicated by the curves in figure 5. Following a depolarization of about 70 mv, the current, which is initially small and outward, reverses rapidly and reaches a maximum in the inward direction at about 0·1 msec at 20°c (or 0·6 msec at 5°c). It subsequently reverses again to approach slowly a large outward intensity at which it is then maintained practically indefinitely. The time course of the current following a depolarization and a hyperpolarization each of 65 mv (for an axon at 3·8°c) is illustrated in figure 6. At smaller depolarizations the transient phase of inward current is less marked ; below 12 mv the current fails to reverse. At larger depolarizations the intensity of this inward current also becomes smaller and beyond a displacement of 110 mv its place is taken by an early phase of outward current. The late phase of maintained outward current is present for all depolarizations, its intensity increasing with the potential change. On hyperpolarization the current is small and is very nearly constant with time.

Further investigation into the ionic current occurring on depolarization has depended on examining the effects produced by changes in the concentration of different ion species in the external medium (Hodgkin and Huxley 1952 a, b, c, d). When the sodium concentration of the solution surrounding the axon is reduced, by replacing this ion with an inert organic cation, the early phase of inward current is abolished and its position in time is occupied by a phase of rapidly developing outward current, similar to that which occurs in sea water when the membrane potential is displaced by more than 110 mv below the resting level. The later slowly developing phase of outward current remains practically unchanged. On the basis of this finding the early phase of current is attributed to a net movement of sodium ions resulting from a large increase in the permeability of the membrane to sodium ions, the direction in which this movement of ions occurs depending on their concentration on the two sides of the membrane and on the membrane potential. With the nerve in sea water the displacement of 110 mv at which the early phase reverses brings the membrane potential close to the level at which

the internal and external concentrations of sodium are in equilibrium, falling short to the extent that the permeability of the membrane is also increased to other ions. For a depolarization of 50 mv which is apparently capable of greatly increasing the permeability to sodium, reduction of the external concentration to below $\frac{1}{10}$ that of sea water results in a reversal from a net inward to a net outward movement of sodium ions. The later outward component of current, which occurs irrespective of the external sodium concentration, is attributed to a delayed increase in permeability to potassium with a resultant net movement of this species of ion.

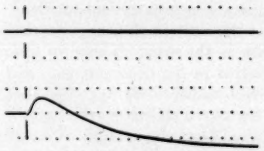


Figure 6. Time course of membrane current elicited by a step of membrane potential. Upper record is obtained on a hyperpolarization of 65 mv, lower record on a depolarization of 65 mv, the displacement of potential from the resting level occurring at the vertical line. Inward current is represented by an upward deflection. Separation of dotted lines gives 10 mA cm^{-2} with the dots 1 msec apart. Records are obtained from a squid axon at 3·8°c. (*After* Hodgkin, Huxley and Katz 1952).

Comparison of the membrane current obtained for a given depolarization at two widely differing concentrations of sodium can be used to separate the components of current carried by sodium and potassium ions. It is assumed that the components due to the two species are independent of each other and that the variation in permeability of the membrane to sodium is a function of membrane potential and time but is independent of current. It is also assumed that the earliest part of the rise in current that follows a sudden depolarization is due entirely to the sodium component. The validity of these assumptions appears to be borne out by the simple manner in which a wide range of observations with different external sodium concentrations can then be made mutually consistent as well as consistent with the approximately known internal concentrations of sodium and potassium. The computed sodium and potassium components of the current following a sudden depolarization can be used to describe the changes in the membrane conductance contributed by these two species of ions. For this purpose the separate components of current are divided by the electrical driving force operating on the relevant ions. This driving force is taken as the difference between the potential at which the current occurs and that at which sodium or potassium ions would be in equilibrium across the membrane. The membrane conductances for sodium, G_{Na}, and potassium, G_K, are thus defined by the equations

$$G_{Na} = I_{Na}/(E - E_{Na})$$

$$G_K = I_K/(E - E_K)$$

where I_{Na} and I_K are the ionic currents and E_{Na} and E_K are the equilibrium potentials. The form of the curves obtained for G_{Na} and G_K is shown in figure 7.

These conductances account for the large currents flowing on depolarization. In addition there will be a comparatively small leakage conductance G_l which is constant for all potentials and gives the linear relation of membrane current and voltage on hyperpolarization. It may also be noted that the conductance for a given ion species will be a function of its concentration on the two sides of the membrane as well as being proportional to the permeability of the membrane. The evaluation of G_{Na} does not require any assumption to be made as to how its magnitude will depend on the concentration, but only that the time course will be unaffected. Deviations would be expected to occur from the above defined

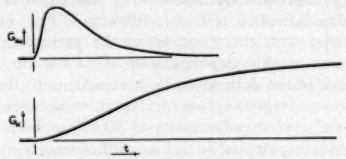


Figure 7. Form of alterations in membrane conductance to sodium (*above*) and potassium (*below*) on a step of depolarization.

relations for the individual ionic conductances when the relative concentrations on the two sides of the membrane are very different, as when the external concentration is reduced to zero. Under these circumstances even though the permeability of the membrane for the particular species of ion were constant, there would still be an appreciable non-linearity between the ionic current and the membrane potential as a consequence of the movement of ions from regions of very different concentrations.

For displacements of potential beyond a depolarization of about 50 mv the maximum values of G_{Na} and G_K are almost constant, each being about $0{\cdot}03\Omega^{-1}$ cm^{-2}. The rates at which G_{Na} and G_K develop on the other hand continue to increase with increasing displacement of potential; beyond a depolarization of 50 mv the maximum rates of rise vary nearly linearly with the potential. Variations in temperature produce a marked alteration in the rates of change of G_{Na} and G_K at a given potential, while only slightly affecting their maximum values. Between 0 and 20°C a 5° increase in temperature compresses the time scale of the curves for G_{Na} and G_K by a factor of about 1·75.

If the membrane potential is suddenly shifted to a new level, at a moment when the conductance towards sodium and potassium ions is raised as a consequence of a previously initiated depolarization, there is a discontinuity in the time course of the current, due apparently to the fact that the membrane conductance does not change abruptly. Thus the current flowing immediately after a displacement of potential from a constant level of depolarization to various other levels can be used to determine the characteristics of the membrane at that moment. The sodium and potassium components of current can be separated owing to the different time courses followed by G_{Na} and G_K, short pulses of preliminary depolarization being used to obtain I_{Na} and long pulses I_K. The principal additional information obtained by this method is that E_K corresponds to a level of potential 15 mv above

the resting potential. The deviation of the resting potential from E_K is attributed to the leakage conductance G_l which would be due to a permeability to chloride ions as well as to components of permeability to sodium and potassium ions which are unaffected by membrane potential. This conductance is liable to great increase on deterioration of the axon. In the normal axon G_l is of about the same magnitude as G_K at the resting potential, i.e. $5 \times 10^{-4} \Omega^{-1}$ cm^{-2}. The level of potential E_l at which the leakage conductance contributes no current will then be at about 15 mv below the resting potential.

A prolonged small depolarization (about 10 mv), while itself producing hardly any increase in G_{Na}, depresses by about 40% the value of G_{Na} reached by a subsequent large depolarization (50 to 100 mv). The effectiveness of this depressant action varies with the duration of the preliminary potential change. On increasing the size of this depolarization, the rate of appearance and the maximum effectiveness of the depression is increased, until for a large depolarization, for which there is a large increase in G_{Na}, it determines the decline in G_{Na}. Thus after a single pulse of depolarization of 50 mv or more, and after G_{Na} has declined, a further shift in potential in the same direction is unable to bring about a second phase of increased G_{Na}.

The qualitative difference between the time courses of G_{Na} and G_K is that the former is transient, while the latter after a prolonged phase of development appears to be maintained indefinitely. These variations in membrane conductance to sodium and potassium may be treated as the result of three processes: a rise of G_{Na}, a decline of G_{Na}, and a rise of G_K, each being a function of membrane potential and time. The manner in which each process varies with time following a shift from any one level of potential (1) to another (2) can be satisfactorily described by a function of the form

$$[m_2 - (m_2 - m_1) \exp(-k_2 t)]^n$$

where m and k are functions of voltage, the former being a measure of the process in the steady state at the indicated voltage and the latter a measure of the rate at which this is approached, while n is a constant. This expression with $n = 4$ can be used directly to give a reasonably good fit of the time course of G_K. To fit G_{Na} the product of two such expressions must be used, one with $n = 3$ for the rise of G_{Na}, and the other with $n = 1$ for its decline. The remaining parameters (m and k for each of the three processes) are obtained as empirical functions of membrane potential by fitting the experimental curves for various displacements of potential. With this information, it is possible to predict the behaviour of G_{Na} and G_K during a voltage change of any time course, by dividing it into a succession of small steps.

The action potential can evidently be accounted for in terms of changes in membrane conductance to sodium and potassium. By virtue of the fact that for a given depolarization the rise in G_{Na} occurs faster than the rise in G_K, the potential approaches close to E_{Na}, the equilibrium potential for sodium, before an increase in G_K together with a decrease in G_{Na}, due to the depression of this component, causes the potential to return rapidly to its resting level. The prolonged hyperpolarization of about 15 mv, which succeeds the rapid phase of the action potential, may be accounted for by the slow decline of G_K at the end of a

large depolarization, causing the membrane potential to be displaced towards E_K. The time course of the action potential has been computed with the use of the above formulation, based on data derived from experiments with constant voltage pulses. Under the condition of zero membrane current, the membrane potential will be determined by the relation

$$C_m\, dE/dt + G_{Na}(E - E_{Na}) + G_K(E - E_K) + G_l(E - E_l) = 0.$$

When E is initially displaced by 12 mv or more below the resting level, the subsequent variation of E with time which this predicts is very close to the observed action potential. This manner of reconstruction of the behaviour of the nerve fibre membrane has also been applied to the action potential propagated along the nerve fibre. In this case the initial part of the action potential, during which current is flowing outward across the membrane, is accounted for, as is also the conduction velocity. A further use to which this procedure has been put is the calculation of the net transfer of sodium and potessium ions across the membrane during a complete action potential by integrating over I_{Na} and I_K. This indicates that in a single action potential there will be a net inward movement of sodium and a net outward movement of potassium, each amounting to about 4×10^{-12} mole cm^{-2}, which is in good agreement with the results of experiments in which changes have been observed in the ionic content of the axoplasm and the surrounding fluid associated with the conduction of impulses (Keynes and Lewis 1951, Shanes 1954).

It may be concluded from the foregoing description that the property of the nerve fibre, to which its characteristic electrical activity is most directly related, is the alteration of the permeability of its membrane towards sodium and potassium ions. Some general remarks may be made about the processes underlying these permeability changes. Firstly, there appears to be a high degree of specificity with respect to the chemical identity of the ions to which these changes apply. This suggests the existence of molecular interaction between the ion traversing the membrane and some constituent of the membrane requiring that a particular fit be obtained. Secondly, the dependence of G_{Na} and G_K on membrane potential indicates that the mechanism responsible for the changes in permeability involves charged bodies or bodies having an electric dipole, the state of which will be influenced by the potential field. Thirdly, the gradual variations in conductance with time following a displacement of potential and, particularly, the continuous nature of the changes in G_{Na} and G_K which occur at the moment when the potential is suddenly shifted make it highly probable that these changes in conductance require some molecular rearrangement of the structure of the membrane. This is further supported by the finding that the rate of change of conductance is highly dependent on the temperature, while the maximum values of G_{Na} and G_K (or the steady state intensities of the underlying processes) are not. While it has been found necessary to assume the independence of I_{Na} and I_K in order to interpret the behaviour of the membrane with different external solutions on the basis of the passive movements of sodium and potassium ions, there is no suitable variable from which it can be inferred that G_{Na} and G_K are also independent. It is possible that there exist side by side in the membrane separate mechanisms for the passage of sodium and potassium. Alternatively, the same mechanism may be used by both species. In this case the structure involved in the passage of these ions may

be pictured as undergoing three alterations, corresponding first to a rise in G_{Na}, then to a decline in G_{Na}, and finally to a rise in G_K, each of the processes occurring with a delay necessitated by molecular rearrangement. The similar maximum values for G_{Na} and G_K obtained following a large depolarization suggests that there may be some limiting stage in the passage of ions which is the same irrespective of which species is moving. Finally, it may be stated that whatever the precise nature of the mechanism, the structure involved in its operation probably constitutes only a small fraction of the fibre membrane, since the rearrangements of this structure during the action potential do not have a significant effect on the membrane capacity, and the conductivity of the membrane, although increasing by a factor of 100, still remains 10^{-5} times that of the surrounding fluid.

§3. Myelinated Nerve Fibre

The nerve fibres of fast conduction in vertebrates differ from those in other animals in that the aqueous internal medium, the axoplasm, is in most places separated from the external medium by a layer of fatty material, called myelin, which accounts for an appreciable fraction of the radius of the axon and which gives vertebrate nerves their characteristic white sheen. From x-ray diffraction and electron microscopy myelin is found to have a laminated fine structure parallel to the surface of the fibre with a repetition interval of approximately 150Å. This interval includes two oppositely orientated lamellae, each composed of fat and protein and about 70Å thick (Robertson, unpublished observations). The arrangement is similar to that existing at the surface of the squid axon, except that the gaps between lamellae are smaller and their number is very much greater. The myelin layer is interrupted at fairly regular intervals along the length of the fibre by narrow gaps, the nodes of Ranvier, where the internal medium is probably separated from the exterior by only a single 70Å membrane.

The larger axons from the nerves in the legs of the frog or toad, the electrical properties of which have been intensively studied, have diameters of 12 to 15μ. The myelin layer in these fibres has a thickness of about 2μ and extends along the fibre for 2 to 2·5 mm between nodes of Ranvier. The internal diameter is reduced from 8 to 10μ in the internodal portion, to about 4μ at the node. The width of the node is less certain, but is probably about 3μ, making the area of membrane exposed at a node about 40μ^2. In the usual experimental procedure, after the nerve is removed from the animal, a central portion several millimetres in length is dissected to leave only a single fibre with a few nodes visible. The fibre is then placed to extend through two or three pools of fluid which serve as wide external electrodes, the low resistivity of the fluid preventing potential gradients occurring within them, and the air gap between pools is bridged by an internodal portion of fibre. A more precise location of the electrical contacts with the fibre is achieved by threading the isolated part of the axon through small holes in a series of insulating sheets, the fluid-filled spaces between them serving as electrodes.

The distinguishing feature of the myelinated fibre is the non-uniform propagation of the nerve impulse, which can be directly attributed to the occurrence of the periodically interrupted layer of myelin. According to the theory of saltatory conduction, the active process which is required for the transmission of the impulse occurs at the nodes of Ranvier, the internodal lengths of fibre behaving as cable

conductors with the myelin as an insulation between the core and the surrounding fluid. Evidence for a separation of the properties of the nodal and internodal regions of the axon comes from studies on the conditions for initiating an impulse in a part of an axon including a small number of nodes (Tasaki 1939, 1940). It is demonstrated that a voltage pulse of a duration of about 1 msec applied between two pools along the length of a fibre causes current to flow between the interior and the exterior of the fibre mainly at the nodes, where alone an impulse can be initiated. This provides the justification for considering, in the first instance, only the currents flowing across the nodes in the interpretation of potential changes recorded between points along the exterior of a fibre during the propagation of an impulse. With the axon placed across two pools, the potential difference recorded between them will be proportional to the current flowing through the layer of fluid adhering to the internodal region joining the pools, which resistance will normally be small compared to the internal resistance and hence will not influence the current. When an impulse is initiated towards one end of the fibre and propagates along the length of fibre joining the pools, a potential wave is recorded which consists of a sharp triangular spike with a duration of about 0·2 msec, the near pool (i.e. the pool on the side of the gap from whence the impulse originates) becoming negative with respect to the other (Tasaki and Takeuchi 1941). This is succeeded by a smaller deflection in the opposite direction continuing for about 1 msec, as indicated by the curve representing longitudinal current in figure 8. When the far pool contains an anaesthetic, e.g. cocaine, which prevents the nodes in it becoming active, the potential wave recorded between the pools has an altered form. Following the initial rapid rise of negativity of the near pool, which is similar to that observed in the previous case, the potential declines relatively slowly at an almost constant rate to about $\frac{1}{3}$ its maximum amplitude, then falling at an increasingly rapid rate and returning to the original level along an S-shaped curve. The duration of the response is about 1 msec at 20°C. It is evident that the potential recorded in this case represents the current flowing away from the anaesthetized nodes in the far pool throughout the period of activity of the nodes in the near pool. In the other case the current in this direction is quickly terminated when the nodes in the far pool become active and the difference in membrane potential between the nodes in the neighbourhood of the gap is reduced. (An appreciable fraction of the current during the first 0·1 msec, however, passes outward through the myelin layer of the internode, as indicated in figure 8.)

Several methods have been employed to determine the amplitude of the membrane potential change occurring at the active node. A rough determination of the size of the action potential can be made from measurements of the current flowing along the fibre in the experiment described above together with information on the internal longitudinal resistance between nodes and the transverse resistance at inactive nodes (Tasaki and Takeuchi 1942). In another method four fluid spaces, separated by insulating partitions through which the fibre penetrates, are arranged along a single internode (Huxley and Stämpfli 1951 a). The nodes on one side of this assembly of electrodes are made inexcitable. A rectangular voltage pulse is applied between one adjacent pair of electrodes, while potential changes, indicative of the current flowing along the fibre, are recorded between another adjacent pair.

When the cathode is nearer to the excitable node than the anode is, the pulse causes current to flow outward at this node which, on attaining a critical amplitude, causes an action potential to be set up there. When the pulse is too weak to excite, it produces a current flowing away from the node, as detected by the potential recorded between the other pair of electrodes. The appearance of the action potential causes this to reverse to one flowing towards the node. The applied voltage is increased until the potential difference recorded at the summit of the active response is the same as before the application of the voltage. In this condition there is no current flowing along the fibre between the nodes on either side of the electrodes, the voltage applied between two of the electrodes being

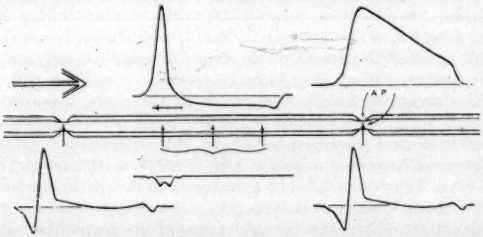


Figure 8. Form of time course of longitudinal current in internodal region (*above, centre*), and currents between inside and outside at each of two nodes (*below, left and right*) and in an internode (*below, centre*). An upward deflection represents a longitudinal current toward the left, and a penetrating current inward. Direction of propagation of impulse is toward the right. An action potential occurring at a node is given above (*right*).

evidently equal and opposite to the potential change developed across the nodal membrane at the summit of the action potential. Thus the applied voltage may be taken to measure the amplitude of the action potential. An extension of the latter procedure employs a feed-back system by which a voltage capable of rapid variation is applied between one pair of electrodes to minimize the potential difference occurring between another pair when an impulse propagates up to the node (Frankenhaeuser 1957). This method permits the complete time course of the action potential to be recorded. In another method potential differences are recorded between two adjacent nodes while the fluid adhering to the internodal region, which is suspended in air, is removed by drying (Tasaki and Frank 1955). As drying proceeds, the potential change recorded in response to an impulse propagated up to this region increases to a limiting amplitude, which is evidently the result of the external longitudinal resistance becoming large compared to the resistance of the rest of the circuit in which current might flow, viz. the interior of the fibre and the inactive nodes. Under these conditions the difference in membrane potential between the active and inactive nodes will be directly recorded. Since on drying the internode the resistance between the recording leads increases to about 60 MΩ, care must be taken to keep the capacity between the leads small to avoid distortion in the recording of rapid changes.

The results of these several methods are in agreement in giving an amplitude of 110 mv for the action potential. The main difference between them is whether

current is or is not flowing across the active nodal membrane. In fact there appears to be no difference in the course of the response beyond the initial 0·1 msec whether it is the membrane current or membrane potential which is observed, there being no current through the capacity of the inactive parts of the fibre beyond the rapid rise of the action potential. The time course of an action potential is indicated in figure 8, its duration being about 1 msec. In those methods in which current is prevented from flowing along the internode, either by applying a potential difference or by increasing the resistance, the nodes on the far side of the internode need not be made inexcitable to record the complete action potential, since there will not be any current to excite them. However, in order to block conduction precaution must be taken to prevent current flowing during the rapid rise of the action potential, either because of a capacity between pools on the two sides of the internode or because of a delay in the feed-back system.

Measurements have also been made on the resting potential at a node by applying a potential difference to abolish current flowing along the internode (Huxley and Stämpfli 1951 a). Reversible electrodes (with silver–silver chloride junctions) are used, and the resting potential is effectively abolished at the nodes on one side of the electrodes by a potassium chloride solution with about the same concentration of potassium as the internal medium. The potential occurring between the axoplasm and the potassium chloride solution is unlikely to be larger than a few millivolts in view of the nearly complete contribution made by potassium to the membrane conductance as is indicated by the effect of varying the external concentration. The resting potential for the node in the normal fluid is determined by this method to be about 70 mv.

The behaviour of the nodal membrane to changes in the external concentrations of potassium and sodium ions is similar to that of the squid axon (Huxley and Stämpfli 1951 b). From a potassium concentration of 5 mmole $l.^{-1}$ (twice that in the normal bathing fluid) to 117 mmole $l.^{-1}$ (presumed to be that in the internal medium) the resting membrane potential varies linearly with the logarithm of the concentration, with a slope of about 50 mv for a ten-fold change in concentration. At concentrations below about 5 mmole $l.^{-1}$ the variation of resting potential is less steep. Changes in sodium concentration have only a slight effect on the resting potential. The action potential on the contrary varies linearly with the logarithm of the sodium concentration ; over a twenty-fold range of concentrations it changes by 58 mv for a ten-fold change at 20°c, in accordance with expectation for a membrane the conductance of which is determined entirely by sodium ions.

The effect of an applied current on the membrane potential of the active node has been investigated. A constant outward current, which depolarizes the resting membrane, causes a reduction in the amplitude of the action potential, while an inward current, which hyperpolarizes the resting membrane, causes an increase, both effects occurring over a limited range (Takeuchi and Tasaki 1943). Examination of the time course of the response reveals that the duration is not altered appreciably and that the rate of slow decline from the summit changes only slightly. The change in amplitude is reflected in the extent of the rapid drop of potential by which the response is terminated. This may be interpreted as indicating that during the slow decline the membrane potential is unaltered by the current because of a low membrane resistance compared with that in the resting condition and

that the termination of the response depends on the potential falling to a critical level. This interpretation is confirmed by the application of alternating current and of short rectangular pulses of current, with which it is furthermore shown that the membrane resistance increases continuously during the slow fall of potential (Tasaki and Freygang 1955).

By the application of a short pulse of inward current at any time during the slow decline of potential, the membrane can be made to revert rapidly to the resting condition (Tasaki 1956). This occurs when the current is above a critical intensity. With weaker currents the membrane potential, which is displaced during the pulse in the direction of the resting level, returns rapidly on the completion of the pulse to continue to follow the normal course of an action potential. For an interval of about 1 msec following a normal action potential, an outward current pulse is incapable of eliciting a second response. Beyond this, up to an interval of about 8 msec, a second action potential can be set up, though of a reduced amplitude and duration which gradually recover as the interval is increased. However, if the action potential is quenched before completing its normal course, another can be set up at any subsequent time by an outward pulse. The response which is elicited within a few tenths of a millisecond of the quenching of a previous one is found to reach an amplitude equal to the height of the earlier action potential when it was cut off, and in fact has the same time course as the final part of the first action potential which was suppressed. With an increasing interval between the two responses the second increases in amplitude and duration in a similar manner to the recovery following a complete action potential. It is concluded from these observations that the membrane potential has two stable positions for zero current. The behaviour of the membrane potential with various currents may be described by the curve shown as a solid line in figure 5. With the membrane initially at rest a constant current can shift the membrane potential to any position on the curve to the right of the minimum, and with the membrane initially active a constant current can shift the potential to any position to the left of the maximum. The region of negative slope, within which the potential cannot be maintained by a constant current, will be passed over rapidly in the transition between the resting and the active condition. The behaviour described by this curve is apparently much less transitory at the node of the myelinated fibre than in the squid axon, making it possible to shift either way between the two stable conditions by the application of current pulses. During the course of an action potential this curve will alter, and the normal action potential will be terminated when the maximum has fallen until it just reaches zero I_1.

The experiments for determining the properties of the active nodal membrane, considered above, all depend on the fact that most of the current that flows between the inside and outside of the fibre crosses at the nodes rather than in the internodal length along which potential differences are observed. The small current that does flow across the myelin has been measured, however. When potentials are recorded between a pool containing only the internodal portion and two other pools shunted together, containing the neighbouring nodes, propagation of an impulse results in the detection of a wave with two sharp peaks separated by about 0·1 msec, the pool containing the internode becoming positive with respect to the others (Tasaki and Takeuchi 1942). When the far lateral pool contains

anaesthetic, only the first such peak appears, the wave being a triangular spike with a duration of about 0·1 msec. A more thorough examination of the distribution of current entering and leaving the fibre is made by passing the axon through a thin insulating partition, the position of which can be varied with respect to the fibre (Huxley and Stämpfli 1949). By this technique the timing of events recorded at different places along a fibre during the propagation of an impulse can be precisely correlated. The individual observations of the potential between the pools on the opposite sides of the partition give the longitudinal current flowing along the portion of internode passing through the partition. Observations made at various positions along an internode show only small differences in wave form. When the position of the partition is moved past a node, the alteration produced in the recorded wave is mainly a displacement in time by about 0·1 msec. This time for each node (or internode) traversed by the impulse agrees with a conduction velocity of 20 to 25 m sec^{-1} as found over a considerable length of fibre. Differences between observations at nearby positions within a single internode give the current crossing the myelin layer, and this may be compared with the current across neighbouring nodes, similarly determined by taking the difference between a pair of observations, one on each side of a node. The time course of these currents is shown in figure 8. The former consists of two brief waves of outward current in agreement with the results of the experiment in which the potential difference between a pool containing an internodal segment and ones containing the neighbouring active nodes is recorded. The current across the nodal membrane consists of an initial outward phase followed immediately by an inward phase of similar intensity, the entire wave lasting about 0·2 msec. Of the two phases of outward current at the internode, the first occurs synchronously with the phase of inward current at the node in the direction from which the impulse is propagating, while the second occurs synchronously with the inward current at the node in the opposite direction. The first wave is also synchronous with the phase of outward current at the far node. These results are consistent with the view that each phase of outward current in the internode occurs during the rising phase of the action potential at the nodes on either side of the internodal length, and the brief time course of each phase indicates that the current flows mainly through the capacity of the myelin layer. The absence of any inward current is an indication that there is no membrane in series with the myelin which becomes active during the propagation of the impulse. An excitable membrane may nevertheless exist in this position without being brought into activity. Because of its low impedance compared with that of the myelin, the potential difference produced across it as a result of activity in the nodes may be insufficient to excite.

From measurements of the current flowing across the myelin during the propagation of an action potential, its resistance and capacity may be determined. Another method for determining these properties of the internode, which has also been applied to an inexcitable node, consists in observing the current flowing outward from a region when action potentials occur synchronously on each side of it (Tasaki 1955). In each method the action potential of nearby nodes provides a voltage pulse which causes current to flow in a passively responding part of the fibre. The internodal region is determined to have a transverse resistance and capacity for unit length of fibre of 30 MΩ cm and 15 pF cm^{-1} with a time constant

of 0·5 msec. From the diameter of the fibre the values for unit area are calculated to be 0·1 MΩ cm² and 0·003 μF cm⁻². With the myelin 2μ thick, including two lamellae within a repetition interval of 150 Å, the total number of lamellae will be about 250. Assuming the resistance and capacity to be the same for each lamella, their specific resistance and capacity will be 500 Ω cm² and 1μF cm⁻². The inactive node is determined to have a transverse resistance and capacity of 40 MΩ and 1·5 pF with a time constant of 0·06 msec. For an estimated area of membrane at the node of 40μ², the specific resistance and capacity will be 15 Ω cm² and 4μF cm⁻².

§ 4. Other Material

From the nerves in the legs of marine crustaceans, single fibres can be isolated over a length of several centimetres with, for the shore crab and lobster respectively, diameters of 30 or 75μ. The cable theory is used to derive membrane properties from potentials measured between various positions along the fibre. For the determination of the electrical characteristics of the fibre in the resting state, a constant current step, up to one half the intensity required to excite, is applied between two distantly separated, narrow electrodes in contact with a fibre suspended in oil, and potential changes are recorded between two other electrodes the positions of which can be varied relative to the current electrodes (Hodgkin and Rushton 1946). According to the cable theory the potential difference occurring between two points along the exterior of the fibre, one of which is near one of the current electrodes, but outside the region between them, and the other a long distance away, will be the final steady potential difference when the near point is at the current electrode multiplied by

$$\tfrac{1}{2}\,e^{-X}\,[1-\mathrm{erf}(X/2\sqrt{T}-\sqrt{T})]-\tfrac{1}{2}\,e^{X}\,[1-\mathrm{erf}(X/2\sqrt{T}+\sqrt{T})]$$

where $T=t/\tau$ and $X=s/\lambda$, t being the time from the start of the current, τ the time constant of the membrane, s the distance of the near point from the current electrode, and λ the length constant of the fibre defined by $\{r_m/(r_i+r_o)\}^{1/2}$. The variation of recorded potentials with position and time is found to agree approximately with this, thus validating the assumptions on the basis of which the theory is applied, viz., that the electrical behaviour of the fibre is determined by the four constants r_i, r_o, r_m and a membrane capacity. For the determination of the first three of these, it is convenient to make three sets of measurements on the potential difference occurring during the application of a steady current. Outside the region between the current electrodes the potential taken with respect to a distant point should vary proportionally to $\exp(-s/\lambda)$. About mid-way between the current electrodes there will be no current crossing the membrane and the potential should vary linearly with distance, the change for unit distance being given by $Ir_ir_o/(r_i+r_o)$, where I is the applied current. Finally, at the current electrode the potential measured with respect to a distant point should be equal to $Ir_o^2\lambda/2(r_i+r_o)$. After having obtained r_m from these observations, the membrane capacity for unit length of fibre can be determined from the time course of potential change at the current electrode (recorded with respect to a distant point) on the application of a current step, which should vary proportionally to $\mathrm{erf}(t/\tau)$. The principal difference between the results of measurements on the crustacean and squid axons

is a higher specific membrane resistance of the former, which ranges up to $8000\Omega\ cm^2$. Deviations from the above theoretical relations occur to some extent around the cathode of the current electrodes, the membrane resistance appearing smaller with increasing outward current (Hodgkin 1947 a). At the anode the membrane resistance is not altered by the current up to a certain level beyond which the resistance diminishes.

An increase in the concentration of potassium in the fluid bathing the axon decreases the membrane resistance (Hodgkin 1947 b). This falls to about one third on treatment with a solution containing three times as much potassium as sea water. When after such treatment the fibre is raised into oil so that only a thin layer of aqueous solution adheres to it, the membrane resistance returns to its normal value in about 10 minutes. This is attributable to the circumstance that the internal potassium concentration is in equilibrium with an external potassium concentration equal to that in sea water; therefore an increase in the external concentration will result in a net entry of these ions. The quantity of external aqueous solution can be estimated from r_o, and from this it is calculated that with a fibre in three times the normal external potassium concentration there occurs a net entry of potassium amounting to about 3×10^{-10} mole $sec^{-1}\ cm^{-2}$. The propagation of a few thousand impulses within a period of a few minutes along a length of axon in oil decreases the membrane resistance to about the same extent as three times the normal potassium concentration (Hodgkin and Huxley 1947). The similarity between the two treatments is stressed by the finding that after the cessation of the nerve impulses the resistance recovers along the same time course as after exposure to a high concentration of potassium, and that the effect of the nerve impulses can be abolished by washing the fibre in a large volume of solution. It is reasonably inferred that the effect of the nerve impulses is a consequence of a release of potassium ions from the interior of the fibre into the small volume of external aqueous solution, and from the number of impulses required to produce a given change it is calculated that the propagation of a single impulse leads to the release of about 2×10^{-12} mole cm^{-2}. This result is confirmed by experiments on the uptake and release of radioactive potassium and on the changes in the amount of potassium in the fluid surrounding a fibre (Keynes 1951, Shanes 1954).

The transfer of a nerve fibre from a large volume of sea water into oil slows the conduction of the impulse, an effect which is attributed to the increase in the external longitudinal resistance (Hodgkin 1939). The influence of external resistance is estimated quantitatively by comparing the effect produced by transferring the fibre to oil from solutions of different electrolyte composition and with different resistivity (Katz 1947). A specific action of the electrolyte on the processes concerned in the production of the action potential is eliminated by considering the ratio of the velocity in oil, where a thin layer of solution adheres to it, to that in a large volume of solution, where effectively $r_o = 0$. The conduction velocity is found to vary inversely proportionally with $(r_1 + r_o)^{1/3}$, in accordance with expectation on the basis that the conduction velocity is affected by the distance to which current will spread along the fibre from an active into an inactive region.

The electrical characteristics of isolated single muscle fibres and of small groups of fibres from the leg muscles of the frog have been examined by means of

narrow longitudinal electrodes for applying current and for recording potentials (Katz 1948). These fibres have diameters of 50 to 100μ and, to avoid rapid deterioration of the surface membrane, must be free from injury along their total length of a few centimetres. This material is distinguished from the large squid and crustacean axons by a larger membrane capacity, amounting to about 6μF cm^{-2}. The difficulty of dissecting single fibres without injuring them has been circumvented by inserting glass capillary electrodes, having a tip diameter of 0·5μ or less and filled with a nearly saturated potassium chloride solution, transversely through the surface membrane and leading from immediately beneath it. This technique is limited, as the fibre is usually seriously damaged around the point of impalement by the contraction which is evoked by a depolarization, including the action potential, and only a single such response can ordinarily be observed in each fibre. Potentials have been recorded between an electrode in the interior and another in the bathing solution, from large numbers of fibres, revealing resting and action potentials that are similar to those found in the squid and crustacean axons, and are affected in essentially the same way by variation in the potassium and sodium concentration of the bathing solution (Ling and Gerard 1949, 1950, Nastuk and Hodgkin 1950). By the insertion of two electrodes into the fibre, one for applying current and another for recording potentials, the effect of a constant current has been observed (Fatt and Katz 1951). From the resultant alterations in the time course of the action potential, involving a greater change of potential on the early part of decline of potential than before or after, it is inferred that the membrane undergoes two phases of increased conductance, corresponding to the rapid rise and fall of the action potential, with a minimum shortly after the summit. Apparently, the increase in sodium conductance which produces the rapid rise has greatly diminished before an increase in potassium conductance occurs to drive the potential towards the resting level more rapidly than would occur if the conductance were that of the resting membrane. Besides fibres behaving in this manner, there occurs in some muscles another type of fibre which does not exhibit any sign of the mechanism for the production of the action potential, i.e. a mechanism for an increase in sodium conductance depending on the membrane potential (Burke and Ginsborg 1956). These fibres are further distinctive in showing a marked non-linearity in the membrane current–voltage relation, the differential resistance increasing on hyperpolarization and decreasing on depolarization. The change occurs with a considerable delay on the displacement of potential from one level to another, and is evidently the result of an alteration in the permeability of the membrane to potassium ions.

The muscle fibres in the leg muscles of crustaceans have been examined by means of the internal electrode and have been found to possess some distinctive properties. It is found that, while the membrane resistance in the resting condition varies widely in different species of animal, the capacity is consistently very high, being about 30μF cm^{-2} (Fatt and Katz 1953 a). With the muscle in the normal bathing medium, which is designed to have a similar ionic composition to that of the animal's blood, many muscle fibres are unable to propagate action potentials, though variations in potential, often oscillatory, occur in the neighbourhood of where the membrane is depolarized by an outward current. These local responses appear to involve a similar change in the electrical properties of the membrane as

occurs in the development of an action potential. They are reduced by the removal of sodium from the bathing medium, but are not completely abolished as long as calcium ions are present. Increasing the concentration of calcium in the bathing medium causes an increase in their size. On the addition of strontium or barium, propagated action potentials can be elicited which are of large amplitude and of very long duration (Fatt and Ginsborg, unpublished observations). All other ions can then be removed from the solution without significantly altering the response. It is inferred that in this material all these species of ions are able to move through the membrane more readily when it is depolarized.

Within the spinal cord of vertebrates there occur the cell-bodies of motoneurones from which arise some of the myelinated fibres that run in the nerves of the leg. The cell-body, or soma, is of an irregular shape, very roughly approximating a sphere with a diameter of about 60μ. From it proceed numerous cylindrical outgrowths. One of these acquires myelin about 50μ from the cell-body and extends outside the spinal cord as the axon. The others which are called dendrites ramify and extend with decreasing diameters for about 0·6 mm in all directions, presenting a surface area several times greater than that of the cell-body. When an electrode is inserted into the cell-body of such a neurone in the spinal cord of the cat, an action potential is recorded in response to the propagation of an impulse from along the axon. This potential rises in two stages, there being a period during its development when the rate of rise decreases or remains constant between periods in which the rate of rise is increasing (Brock, Coombs and Eccles 1953). The rate of rise of potential during both phases of the action potential is similarly decreased by an outward current applied at the soma, which decreases the membrane potential there from its resting level (Coombs, Eccles and Fatt 1955 a). From the results of recording potential changes at various positions in the external medium when a single motoneurone is activated by the propagation of an impulse along its axon, the currents flowing between different parts of the motoneurone during the two phases in the development of the action potential are determined (Fatt 1957 a). It is concluded from this study that the first phase involves the activity of a considerable fraction of the massive central part of the motoneurone, which may be identified with the cell-body. The second phase, which commences 0·1 to 0·2 msec from the first, involves the activity of the dendrites along which the activity propagates away from the cell-body at a rate of about 1 m sec^{-1}. This may be compared with the rate of propagation in the myelinated axon of the same motoneurone which at the body temperature of 38°c is about 100 m sec^{-1}.

A quite different process from that which is concerned in the propagation of the impulse in nerve and muscle fibres operates in the transmission of excitation from nerve to muscle. There is evidence to indicate that in the vertebrate the arrival of an impulse at the nerve terminals in contact with the muscle fibre causes the release of a specific substance, acetylcholine, which on reacting with a specially sensitive region of the muscle fibre surface alters its properties, causing a lowering of membrane potential which may lead to the production of an action potential. The local alteration in membrane properties occurring in the muscle fibre is fundamentally different both from the effect of an applied current and from the processes involved in the production of an action potential. On the arrival of

the nerve impulse, the membrane potential is displaced towards a level near that of zero membrane potential (Fatt and Katz 1951). Thus it is observed that, while the resting membrane is depolarized, the peak of the action potential, which extends normally well past the level of zero membrane potential is lowered. It is concluded that the action of acetylcholine is to increase the permeability of the membrane indiscriminately to all ions present on both sides.

The crustacean muscle fibres are contacted by two types of nerve fibre, one of which is capable of causing an alteration in the electrical properties similar to that occurring in the vertebrate muscle fibre, though acetylcholine does not appear to be involved, while the other type of fibre exerts an inhibitory action (Fatt and Katz 1953 b). If an impulse in such a fibre arrives at the junctional region in suitable time relation to an impulse in the former type of fibre, the excitatory one, it suppresses the potential change which would otherwise be elicited. An impulse in the inhibitory fibre by itself does not ordinarily cause any change of membrane potential from its resting level, but it causes an increase in conductance, this conductance being in series with an e.m.f. equal to the resting potential. Although this will have the effect of diminishing any change of potential from the resting level, such as is produced by the excitatory nerve fibre, the observed effect of an inhibitory impulse cannot be accounted for entirely by this action, and it is necessary to postulate a competition between two hypothetical substances released from these two types of nerve fibres at a stage preceding the alteration of the electrical properties of the membrane.

In the vertebrate many nerve fibres carrying impulses from various parts of the nervous system make contact with each motoneurone in the region of the cell-body and dendrites. Some of these have an inhibitory and some an excitatory action. The excitatory fibres have an effect here similar to that in other material, i.e. they cause a displacement of potential towards the level of zero membrane potential, while the inhibitory fibres cause an increase in conductance which shifts the membrane potential towards a level about 10 mv above the resting potential (Coombs, Eccles and Fatt 1955 b, c). By the injection of ions into the motoneurone from the capillary that is otherwise used to record potentials or to apply current, it is demonstrated that inhibitory action involves an increase in membrane permeability at least to chloride ions and possibly to potassium, but not to sodium. In the normal behaviour of the motoneurone the possibility of an impulse being set up at a given moment depends on the relative intensity of excitatory and inhibitory actions, an impulse being initiated when the membrane potential is decreased by 10 mv. Geometrical factors will be involved to the extent that excitatory and inhibitory actions will be exerted over the large surface area of the dendrites as well as the cell-body, but the impulse is normally initiated in the region of the cell-body because of the displacement required to set up an action potential being less there than in the dendrites (Fatt 1957 b).

References

Bear, R. S., and Schmitt, F. O., 1939, *J. Cell. Comp. Physiol.*, **14**, 205.
Brock, L. G., Coombs, J. S., and Eccles, J. C., 1953, *J. Physiol.*, **122**, 429.
Burke, W., and Ginsborg, B. L., 1956, *J. Physiol.*, **132**, 586.
Cole, K. S., 1941, *J. Gen. Physiol.*, **25**, 29.

COLE, K. S., and BAKER, R. F., 1941 a, *J. Gen. Physiol.*, **24**, 535 ; 1941 b, *Ibid.*, **24**, 771.
COLE, K. S., and CURTIS, H. J., 1938, *Nature, Lond.*, **142**, 209 ; 1939, *J. Gen. Physiol.*, **22**, 649 ; 1941, *Ibid.*, **24**, 551.
COLE, K. S., and HODGKIN, A. L., 1939, *J. Gen. Physiol.*, **22**, 671.
COOMBS, J. S., ECCLES, J. C., and FATT, P., 1955 a, *J. Physiol.*, **130**, 291 ; 1955 b, *Ibid.*, **130**, 326 ; 1955 c, *Ibid.*, **130**, 374.
CREMER, M., 1909, *Nagels Handbuch der Physiologie*, **4**, 793.
CURTIS, H. J., and COLE, K. S., 1938, *J. Gen. Physiol.*, **21**, 757 ; 1940, *J. Cell. Comp. Physiol.*, **15**, 147 ; 1942, *Ibid.*, **19**, 135.
FATT, P., 1957 a, *J. Neurophysiol.*, **20**, 27 ; 1957 b, *Ibid.*, **20**, 61.
FATT, P., and KATZ, B., 1951, *J. Physiol.*, **115**, 320 ; 1953 a, *Ibid.*, **120**, 171 ; 1953 b, *Ibid.*, **121**, 374.
FRANKENHAEUSER, B., 1957, *J. Physiol.*, **135**, 550.
GEREN, B. B., and SCHMITT, F. O., 1954, *Proc. Nat. Acad. Sci., Wash.*, **40**, 863.
GOLDMAN, D. E., 1943, *J. Gen. Physiol.*, **22**, 37.
HODGKIN, A. L., 1939, *J. Physiol.*, **94**, 560 ; 1947 a, *Ibid.*, **106**, 305 ; 1947 b, *Ibid.*, **106**, 319.
HODGKIN, A. L., and HUXLEY, A. F., 1939, *Nature, Lond.*, **144**, 710 ; 1945, *J. Physiol.*, **104**, 176 ; 1947, *Ibid.*, **106**, 341 ; 1952 a, *Ibid.*, **116**, 449 ; 1952 b, *Ibid.*, **116**, 473 ; 1952 c, *Ibid.*, **116**, 497 ; 1952 d, *Ibid.*, **117**, 500.
HODGKIN, A. L., HUXLEY, A. F., and KATZ, B., 1952, *J. Physiol.*, **116**, 424.
HODGKIN, A. L., and KATZ, B., 1949 a, *J. Physiol.*, **108**, 37 ; 1949 b, *Ibid.*, **109**, 240.
HODGKIN, A. L., and RUSHTON, W. A. H., 1946, *Proc. Roy. Soc.* B, **133**, 444.
HUXLEY, A. F., and STÄMPFLI, R., 1949, *J. Physiol.*, **108**, 315 ; 1951 a, *Ibid.*, **112**, 476 ; 1951 b, *Ibid.*, **112**, 496.
KATZ, B., 1947, *J. Physiol.*, **106**, 411 ; 1948, *Proc. Roy. Soc.* B, **135**, 506.
KEYNES, R. D., 1951, *J. Physiol.*, **113**, 99.
KEYNES, R. D., and LEWIS, P. R., 1951, *J. Physiol.*, **114**, 151.
LING, G., and GERARD, R. W., 1949, *J. Cell. Comp. Physiol.*, **34**, 383 ; 1950, *Nature, Lond.*, **165**, 113.
MARMONT, G., 1949, *J. Cell. Comp. Physiol.*, **34**, 351.
NASTUK, W. L., and HODGKIN, A. L., 1950, *J. Cell. Comp. Physiol.*, **35**, 39.
SCHMITT, F. O., and SCHMITT, O. H., 1940, *J. Physiol.*, **98**, 26.
SHANES, A. M., 1954, *Amer. J. Physiol.*, **177**, 377.
STEINBACH, H. B., 1941, *J. Cell. Comp. Physiol.*, **17**, 57.
STEINBACH, H. B., and SPIEGELMAN, S., 1943, *J. Cell. Comp. Physiol.*, **22**, 187.
TAKEUCHI, T., and TASAKI, I., 1943, *Pflügers Arch. ges. Physiol.*, **246**, 32.
TASAKI, I., 1939, *Amer. J. Physiol.*, **127**, 211 ; 1940, *Pflügers Arch. ges. Physiol.*, **244**, 125 ; 1955, *Amer. J. Physiol.*, **181**, 639 ; 1956, *J. Gen. Physiol.*, **39**, 377.
TASAKI, I., and FRANK, K., 1955, *Amer. J. Physiol.*, **182**, 572.
TASAKI, I., and FREYGANG, W. H., 1955, *J. Gen. Physiol.*, **39**, 211.
TASAKI, I., and TAKEUCHI, T., 1941, *Pflügers Arch. ges. Physiol.*, **244**, 696 ; 1942, *Ibid.*, **245**, 764.
WEBB, D. A., and YOUNG, J. Z., 1940, *J. Physiol.*, **98**, 299.

THE THEORY OF COLLECTIVE EXCITATIONS IN NUCLEI

By D. M. BRINK
Clarendon Laboratory, Oxford

CONTENTS

	PAGE
§ 1. Introduction	145
§ 2. Individual particle and collective aspects	145
2.1. Hartree–Fock approximation	145
2.2. Configuration mixing and collective effects	146
2.3. Surface vibrations	147
§ 3. Nuclear deformations	148
3.1. The oscillator model	148
3.2. Intrinsic electric and magnetic moments	151
§ 4. Nuclear rotations	151
4.1. The symmetric top in quantum mechanics	151
4.2. Dynamics of nuclear deformations	152
4.3. Rotational energy levels	155
4.4. Moments of inertia	155
4.5. Electric and magnetic moments	156
4.6. Electromagnetic transition probabilities	158
4.6.1. Collective transitions	159
4.6.2. Particle transitions	159
§ 5. Nuclear vibrations	161
§ 6. Recent theoretical developments	161
Acknowledgments	163
References	163

Abstract. In some regions of the periodic table nuclei show an energy spectrum characteristic of a rotator and appear to have a non-spherical equilibrium shape. Such nuclei have been described successfully by a unified model which exhibits both collective and individual particle aspects of nuclear structure. The model ascribes wave functions to the low nuclear states which are products of collective wave functions, representing the rotation of the nuclear deformation, with wave functions representing the intrinsic motion of nucleons inside the nucleus.

Sequences of levels showing the rotational spectrum correspond to states with the same intrinsic wave function, but with different collective motions. The rotational wave functions are represented by rotation matrices, and the collective motion is closely analogous to the quantum mechanical rotations of a symmetric top with moment of inertia about half that of an equivalent rigid body.

The intrinsic state is approximated by a Hartree–Fock type wave function, with the individual particle orbitals defined by a deformed potential determined by the nuclear shape. Magnetic dipole and electric quadrupole moments are given by the corresponding intrinsic moments modified by a projection factor depending upon the mode of collective rotation. Hence the intrinsic deformation of the nuclear charge gives rise to large electric quadrupole moments. Within a rotational sequence, electromagnetic transition probabilities depend on the intrinsic electric or magnetic moments and on the rotational modes of the initial and final states. Between rotational sequences, where the intrinsic particle state changes, the transition probabilities depend on an 'intrinsic transition probability' modified by parameters determined uniquely by the collective motion.

Other nuclei show an energy spectrum of a type which can be associated with a vibration of the nuclear surface.

§ 1. Introduction

Considerable progress has been made in the theoretical description of the nucleus in recent years. There is strong evidence that the nucleons in a nucleus can be considered as occupying states characteristic of independent particle motion in an average nuclear field. This aspect of nuclear structure has been developed in the shell model. On the other hand, collective views of nuclear structure, comparing the nucleus to a liquid drop, were proposed early in the history of nuclear physics in order to describe the phenomenon of nuclear fission.

The occurrence of nuclear quadrupole moments with values many times those associated with a single particle orbit has led to the view that in some regions of the periodic table nuclei may be strongly deformed. Further, many of these nuclei have an energy spectrum characteristic of a rotator, i.e. energy levels are spaced according to the law

$$E_I = \frac{\hbar^2}{2\Im} I(I+1)$$

where I is the angular momentum of the excited state in question. It is believed that the observed spectrum arises from a collective motion of the nuclear matter associated with a rotation of the deformation. Nuclei with these rotational properties are found in the neighbourhood of atomic weights $A \simeq 20$, $150 < A < 190$ and $A > 220$. In other regions of the periodic table odd nuclei have quadrupole moments consistent with large deformations of the nuclear charge, while neighbouring even–even nuclei have an energy spectrum of a type which can be associated with surface vibrations of a liquid drop. In fact, these characteristic vibration spectra are found in most parts of the periodic table away from closed shells and outside the rotational regions.

This review contains an analysis of the collective and individual particle aspects of the unified model first proposed by Bohr (1952) and Bohr and Mottelson (1953, 1954) and aims at illustrating the methods of calculation used in developing the model and obtaining results from it. No attempt is made to discuss the vast amount of experimental material now available which is relevant to the model. For this readers are referred to Chapter V of the article by Alder, Bohr, Huus, Mottelson and Winther (1956).

§ 2. Individual Particle and Collective Aspects

2.1. *Hartree–Fock Approximation*

A representation in terms of an approximation in which interactions between the electrons of an atom are replaced by an average field is remarkably useful as a starting-point for the study of atomic structure and atomic spectra. The consequence of ignoring the fluctuations in the interelectronic interactions, which is implicit in the self-consistent field method, is that the energy states of an atom are represented by wave functions which are Slater determinants of one-electron wave functions, each corresponding to a definite single particle state in the self-consistent field.

The extension of this concept to the nucleus provides the basis of the nuclear shell model, which has been so successful in the analysis of nuclear spins and

nuclear magnetic moments, and which has provided the starting-point for an extensive study of the structure of light nuclei. The application of the self-consistent field method to nuclei is much more open to question than it is in the atomic case, since, in contrast to the long-range Coulomb potential, the nuclear forces are of short range, and indications are that they are very singular with strong repulsive cores. This last fact would make complete nonsense of any application of the Hartree method. It has been demonstrated, however, that these singular potentials do not destroy the independent particle picture of the nucleus (Brueckner 1955, Bethe 1956); and it is probable that, though the repulsive cores may be of considerable importance in the study of nuclear saturation, they play no important part in questions of nuclear spetcra and in particular of the existence of collective rotational and vibrational modes. It is assumed, therefore, that some effective internuclear potential can be chosen which reproduces at least the general features, and to some extent the details, of nuclear spectra, and yet is sufficiently non-singular to give some validity to the Hartree–Fock self-consistent field approximation.

In the shell model the self-consistent field is spherical and has the following qualitative features. As a result of the constant density of nuclear matter, the self-consistent potential is almost constant inside the nucleus and falls rather sharply to zero at the nuclear surface. The effective forces may possess some exchange character, giving the potential a velocity dependence (Van Vleck 1935) which may be represented approximately by an effective nucleon mass (Weisskopf 1957) of about one half the real nucleon mass. In addition there is a strong spin–orbit coupling.

Away from closed shells the Hartree wave functions will not in general be a good approximation to the total nuclear wave function because of the degeneracy of one-nucleon states in the spherical field. Taking the nuclear interactions to a better approximation the wave function would be a linear combination of Slater determinants. The goodness of the shell model wave function depends upon the magnitude of the fluctuations in the average field and on whether there are other configurations with nearly the same energy, for then the real wave function may contain appreciable admixtures of several configurations.

2.2. *Configuration Mixing and Collective Effects*

In cases where a number of configurations are almost degenerate in the shell model configuration mixing may be a manifestation of some collective motion in the nucleus. Consider as an example the excited state formed on absorption of photons by nuclei in electric dipole transitions. The state in question corresponds to the 'giant resonance' in the photo-nuclear cross section occurring in all nuclei at energies which vary from 20 Mev in light nuclei to 13 Mev in heavy nuclei. If we approximate the nucleus by an oscillator shell model in which the protons and neutrons move independently in an oscillator potential, then all excited states of this model nucleus will have energies which are multiples of the oscillator energy $\hbar\omega$. Photons of energy $\hbar\omega$ can be absorbed in dipole transitions and the state formed corresponds to the 'giant resonance' state observed in actual nuclei. On absorption of a photon of energy $\hbar\omega$ single protons and neutrons in the highest filled shells can be excited. A number of different single

particle transitions are possible, but on dipole absorption a particular coherent combination (i.e. a certain configuration mixing) of these single-particle excited states is produced which corresponds to a collective mode discussed by Goldhaber and Teller (1948). Pictorially this mode is a dipole vibration of all the protons in the nucleus against all the neutrons.

The example discussed above is a case in which the configuration mixing associated with the collective mode is a result of the method of excitation—the absorption of a dipole photon produces a dipole type vibration. On the other hand the collective effects of nuclear rotations and surface vibrations are a consequence of the nuclear forces. As will be seen in the following sections the forces are such that away from closed shells a spherical nucleus with its resultant high degeneracies is unstable (Jahn–Teller effect 1937) and a deformed ellipsoidal shape is favoured. Attempts at analysis of nuclear rotations and vibrations in terms of configuration mixing of shell model wave functions have in general been unsuccessful, though recently some advances have been made in nuclei with atomic weight around $A = 20$ (Elliot 1957).

2.3. *Surface Vibrations*

It was pointed out early in the history of nuclear structure by Bohr and Kalckar (1937) that a relative instability of the nuclear shape should imply the existence of low frequency shape oscillations of the nucleus with corresponding low energy excited states. On this basis the analysis of low energy properties of a nucleus would require investigation of the interplay between the collective shape oscillalations and the motion of particles inside the nucleus (Bohr 1952).

If these shape oscillations have a frequency small compared with frequencies associated with single particle motions in the nucleus the adiabatic approximation may provide a valid approach to the problem. A deformed nucleus with shape specified by parameters α would have energy states $E_i(\alpha)$ depending upon α. If now the nucleus makes relatively slow variations in shape the internal state of the nucleus should not be changed (adiabatic approximation) though there will be additional kinetic energy associated with the variations of the shape parameters with time. Thus the total energy of the system which could be taken as the Hamiltonian for the collective motion would have the form

$$E_i = E_i(\alpha) + \tfrac{1}{2}B_i(\alpha)\,\dot{\alpha}^2 \qquad \ldots\ldots(3.1)$$

to the lowest order in α. The wave functions of such a system in the adiabatic approximation has the form

$$\Psi_i = \phi_{i\text{vib}}(\alpha)\,\chi_{i\text{int}}(x, \alpha)$$

where $\phi_{i\text{vib}}(\alpha)$ is the collective wave function and $\chi_{i\text{int}}(x, \alpha)$ is an internal wave function corresponding to a deformation α and energy $E_i(\alpha)$.

A study of the collective motions of the nucleus by this method involves an investigation of the variation of $E_i(\alpha)$ and $B_i(\alpha)$. Originally these parameters were estimated from an irrotational liquid drop model (Bohr 1952), the energies $E_i(\alpha)$ arising from the surface tension and $\frac{1}{2}B_i(\alpha)\,\dot{\alpha}^2$ coming from the collective flow of liquid inside the vibrating drop. These estimates gave little correlation with experiments, however, and it is now recognized that the energies in (3.1)

are strongly dependent on the internal particle structure of the nucleus. Near closed shells the nucleus is very stable against deformations. If, however, we move away from closed shells the nucleus becomes much softer against deformations at first and then tends to become polarized with a deformation large compared with the amplitude of surface vibrations. Thus in the intermediate neighbourhood of closed shells collective vibrations have frequencies of the order of the single-particle excitations in the shell model and the adiabatic approximation has no validity. Away from closed shells the collective vibrations have a much lower frequency and when large deformations are produced due to the polarizing effect of the internal nucleon motion, rotational modes become important.

§ 3. Nuclear Deformations

In the above discussion of the shell model in terms of a Hartree–Fock type approximation the self-consistent field was assumed to be isotropic with the result that in nuclei far from closed shells the shell model states are degenerate. It is known, however, that many nuclei in these regions are deformed as evidenced by the existence of large quadrupole moments. Rainwater (1950) answered this question by showing that between closed shells the last nucleons tended to polarize the nucleus. In other words, if the restriction of isotropy on the Hartree potential is relaxed, a lower self-consistent energy can be obtained. The wave functions are then determinants of single-particle wave functions in a deformed potential. Using this deformed field Hartree–Fock method the spherical symmetry originally present in the problem is lost, and the wave functions obtained are inadequate in that they do not correspond to a definite total angular momentum, i.e. the angular momentum of the model is not conserved whereas the angular momentum of the nuclear system is. Nevertheless, the deformed wave functions provide the ' best ' Hartree–Fock solution in the sense that they give the lowest energy.

As in the case of the shell model, we do not know enough about the exact nature of nuclear forces to solve the Hartree problem, but we can use the same qualitative arguments as are used in the shell model to obtain information about the self-consistent field. Nuclear matter has approximately constant density, hence we assume that the nuclear volume is preserved on deformation, and that the potential inside the nucleus is approximately constant with the same depth as in the spherical shell model. The usual spin–orbit coupling is still present. This deformation problem has been investigated (Nilsson 1955, Moszkowski 1955, Gottfried 1956) by choosing a convenient potential consistent with the above requirements, solving for the particle motion, then varying the shape of the potential to give minimum energy.

3.1. *The Oscillator Model*

The oscillator model as discussed by Gallone and Salvetti (1953) and by Inglis (1952) provides a simple example to illustrate the above remarks. Suppose the nuclear forces are such as to make the oscillator potential a suitable self-consistent potential. In the deformed nucleus it would have the form

$$V_0(x,y,z) = \tfrac{1}{2}M(\omega_1^2 x^2 + \omega_2^2 y^2 + \omega_3^2 z^2) \qquad \ldots\ldots(3.1)$$

with the restriction of constant density (approximately equivalent to a condition

of constant volume of the potential) $\omega_1\omega_2\omega_3 = \text{constant}$. Confining our discussion to the case of cylindrical symmetry we can introduce a single deformation parameter ϵ

$$\omega_1 = \omega_2 = \omega(\epsilon)(1+\tfrac{1}{3}\epsilon),$$
$$\omega_3 = \omega(\epsilon)(1-\tfrac{2}{3}\epsilon). \qquad \ldots\ldots(3.2)$$

The restriction of constant density determines the deformation dependence of $\omega(\epsilon)$ to be

$$\omega(\epsilon) = \omega_0(1-\tfrac{1}{3}\epsilon^2-\tfrac{2}{27}\epsilon^3)^{-1/3}, \qquad \ldots\ldots(3.3)$$

where ω_0 is the underformed oscillator frequency. The parameter ϵ is related to the quantity β used in papers by Bohr and Mottelson in the first order by

$$\epsilon \simeq \frac{3}{2}\left(\frac{5}{4\pi}\right)^{1/2}\beta \simeq 0{\cdot}95\beta.$$

The Hamiltonian of the anisotropic oscillator separates in rectangular coordinates parallel to the principal axes of deformation, hence the wave function for a particle moving in this potential can be written as a product $|n_1\rangle|n_2\rangle|n_3\rangle$ where n_1, n_2 and n_3 are the oscillator quantum numbers for the motion in the x, y and z directions respectively. The particle energy corresponding to the above state is

$$E(n_1,n_2,n_3) = \hbar[(n_1+\tfrac{1}{2})\omega_1+(n_2+\tfrac{1}{2})\omega_2+(n_3+\tfrac{1}{2})\omega_3]$$
$$= \hbar\omega(\epsilon)[N+\tfrac{3}{2}+(n_1+n_2-2n_3)\epsilon] \qquad \ldots\ldots(3.4)$$

with $N = n_1+n_2+n_3$ as the principal quantum number of the oscillator shell model. The energy of a particle depends only on the principal quantum number N in the spherical case, but in the deformed shell model there is an additional term depending on the deformation and on the individual quantum numbers n_1, n_2 and n_3. Provided that the deformation is not so large that levels from neighbouring oscillator shells overlap, the single-particle states fill in a characteristic order for prolate deformation ($\epsilon > 0$) and oblate deformation ($\epsilon < 0$). In a shell with principal quantum number N the lowest state in the shell is $n_3 = N, n_1 = n_2 = 0$ for a prolate deformation, and the levels fill in order of decreasing n_3. For an oblate or negative deformation the order of levels is opposite and the lowest group of states in the shell is $n_3 = 0, n_1+n_2 = N$.

The total energy of the shell model nucleus is the sum of the single-particle energies (3.4). In a closed shell nucleus the term in the energy proportional to ϵ drops out summing over occupied states and, if the Nth shell is filled, the total energy of the model is

$$E = \tfrac{1}{3}\hbar\omega(\epsilon)(N+1)(N+3)(N+2)^3.$$

The deformation dependence lies in the factor $\omega(\epsilon)$ which has its minimum value at zero deformation and increases as ϵ increases. Thus for a closed shell nucleus the spherical shape is stable. A nucleus with an unfilled shell will have an energy of the form $E = \hbar\omega(\epsilon)(a+\epsilon b)$ where the constants a and b are determined from (3.4) by summing over the occupied levels. The equilibrium shape of the nucleus is determined minimizing with respect to ϵ and in general there will be a resultant deformation. As b increases relative to a so the deformation increases and the total energy decreases. Both the oblate and prolate deformations have a lower energy than the spherical shape. It turns out, however, by comparing

the parameter b for oblate and prolate deformations that prolate deformations are favoured in the first half of the shell while oblate deformations have a lower energy in the second half.

The high degeneracies present in the oscillator potential make it an unrealistic representation of the nuclear field, but it can be improved considerably by making the following modification :

$$V = V_0(x, y, z) + C\mathbf{l}.\mathbf{s} + D\mathbf{l}^2 \qquad \ldots\ldots(3.5)$$

where $V_0(x, y, z)$ is the deformed oscillator potential (3.1) and $C\mathbf{l}.\mathbf{s}$ is the usual spin–orbit coupling. The $\mathbf{l}^2$ term corrects the oscillator field by depressing the high orbital angular momentum states ; hence it incorporates some of the features of a square well potential. A suitable choice of C and D can reproduce the known level sequence of the spherical shell model.

Solutions for the motion of a particle in a deformed potential of the form (3.5) no longer have the particle orbital angular momentum and the total particle angular momentum as constants of the motion. Only the z component of the angular momentum $j_z = l_z + s_z$ commutes with the Hamiltonian and has eigenvalues denoted by Ω. For a potential of the form (3.5) which is reflection invariant, parity remains a good quantum number. If the deformation is small, l and j remain approximate quantum numbers.

Large deformations reduce the effect of the spin–orbit coupling and though l and j are no longer good quantum numbers the components l_z and s_z of $\mathbf{l}$ and $\mathbf{s}$ become approximately such with eigenvalues denoted by Λ and Σ.

The problem of particle motion in a potential of the form (3.5) has been discussed by Nilsson (1955) and Gottfried (1956). They expand the exact solutions in terms of spherical oscillator ($\epsilon = 0$) solutions using a representation with l^2, l_z and s_z diagonal and tabulate the coefficients of the expansion for several values of ϵ. Also they give the single particle energies and order of levels as a function of the deformation.

In the deformed oscillator model prolate deformations are favoured at the beginning of the shell and oblate deformations at the end. The situation is altered with the potential (3.5) and it appears that prolate deformations are favoured throughout. In calculating the total energy of the nucleus the correct expression is $E = \Sigma T_i + \frac{1}{2}\Sigma V_i$ where the summations are over all the particles and the T are particle kinetic energies. The factor $\frac{1}{2}$ is present because the self-consistent potential (3.5) arises from two-body nucleon–nucleon interactions, so that V is exactly twice the total potential.

Each level of the deformed oscillator shell model is twofold degenerate : it can be occupied by a nucleon with $j_z = \pm\Omega$. Any nucleus in its lowest state has its levels filled in pairs in order of increasing energy. Thus an even–even nucleus has the z component of the total particle angular momentum J_z equal to zero, and an odd nucleus has $J_z = \pm\Omega$ where Ω is the value of j_z for the last nucleon. An odd nucleus has two distinct possibilities $J_z = \pm(\Omega_p + \Omega_n)$ with Ω_p and Ω_n as the axial components of the angular momentum of the odd proton and odd neutron respectively. Resultant nuclear forces determine which of these possibilities has the lowest energy.

3.2. *Intrinsic Electric and Magnetic Moments*

The wave functions given by Nilsson (1955) and Gottfried (1956) can be used to predict various properties of the deformed nuclear system. The nucleons moving in the distorted field have an associated distorted charge distribution, hence the characteristically large intrinsic quadrupole moment,

$$Q_0 = \langle |\sum_1^Z (3\cos^2\theta_i - 1)| \rangle$$
$$= 0{\cdot}8ZR_Z^2(\epsilon + \tfrac{1}{2}\epsilon^2),$$

as calculated from the distorted oscillator wave functions (Nilsson 1955).

All nucleons of an even–even nucleus are paired off, hence there is no resultant intrinsic magnetic moment. The entire magnetic moment of an odd nucleus is due to the last odd nucleon for which the magnetic moment operator has the form

$$\boldsymbol{\mu} = g_s\mathbf{s} + g_l\mathbf{l}$$

in units of the nuclear magneton. The coefficients g_s and g_l are the spin and orbital g-factors. For free nucleons one has $g_s = (\begin{smallmatrix}5{\cdot}585\\-3{\cdot}826\end{smallmatrix})$ and $g_l = (\begin{smallmatrix}1\\0\end{smallmatrix})$ for protons and neutrons respectively. The intrinsic magnetic moment of the nucleus is the expectation value of the axial component of this operator between the deformed shell model states:

$$\mu_{\text{int}} = g_s\langle s_z\rangle + g_l\langle l_z\rangle.$$

Since $s_z + l_z = j_z$, this can be written as

$$\mu_{\text{int}} = g_l\,\Omega + (g_s - g_l)\langle s_z\rangle.$$

In the strong coupling limit s_z is an approximate constant of the motion with eigenvalue Σ, and the intrinsic moment takes the simple form

$$\mu_{\text{int}} = g_l\,\Omega + (g_s - g_l)\Sigma.$$

The spin component Σ can take values $\pm\frac{1}{2}$, while in intermediate coupling cases the expectation value of the spin $\langle s_z\rangle$ will lie between these limits. Thus the expression (3.8) gives limiting values of the intrinsic moment analogous to the Schmidt limits in the j–j coupling shell model. For later work it is useful to introduce an intrinsic g-factor g_Ω defined by $\mu_{\text{int}} = \Omega g_\Omega$.

§4. Nuclear Rotations

4.1. *The Symmetric Top in Quantum Mechanics*

Consider an axially symmetric rigid body, which can rotate freely about its centre of mass. The kinetic energy of the rigid body is

$$T = \tfrac{1}{2}\mathfrak{I}_1(w_1^2 + w_2^2) + \tfrac{1}{2}\mathfrak{I}_3 w_3^2$$

where w_1, w_2, w_3 are the angular velocity components referred to a set of principal axes fixed in the body, and $\mathfrak{I}_1$, $\mathfrak{I}_3$ are the principal moments of inertia. The axis '3' is the symmetry axis. The components of the rigid body angular momentum referred to the same axes are $\hbar I_1 = \mathfrak{I}_1 w_1$, $\hbar I_2 = \mathfrak{I}_1 w_2$, and $\hbar I_3 = \mathfrak{I}_3 w_3$. In the

absence of external torques the total kinetic energy is equal to the Hamiltonian H of the rigid body and in terms of the angular momenta it takes the form

$$H = \frac{\hbar^2}{2\Im_1}(I_1^2+I_2^2)+\frac{\hbar^2}{2\Im_3}I_3^2 = \frac{\hbar^2}{2\Im_1}(I^2-I_3^2)+\frac{\hbar^2}{2\Im_3}I_3^2. \quad \ldots\ldots(4.1)$$

The angular momentum of the system is conserved, hence I^2 and I_z commute with the Hamiltonian. Also by inspection of (4.1) it is seen that the component of the angular momentum of the top along its symmetry axis is conserved. In fact, I^2, I_z and I_3 form a set of commuting operators with corresponding eigenvalues I, M and K respectively. Substituting in (4.1) we see the energy of the state with these quantum numbers takes the form

$$E = \frac{\hbar^2}{2\Im_1}I(I+1)+\left(\frac{\hbar^2}{2\Im_3}-\frac{\hbar^2}{2\Im_1}\right)K^2. \quad \ldots\ldots(4.2)$$

It may be shown that the eigenfunctions of (4.1) corresponding to the energy (4.2) are the rotation matrices $\mathfrak{D}_{MK}^{I}(\theta_i)$. The arguments of the wave function are the Eulerian angles giving the orientation of the set of principal axes fixed in the body relative to axes fixed in space. Given a particular K the total angular momentum I can take values

$$I = |K|, |K+1|, \ldots . \quad \ldots\ldots(4.3)$$

If $\Im_3 \ll \Im_1$ (a long thin top) the levels corresponding to different values of K are widely spaced compared with the spacing of levels with the same K and different I. For each value of K there is a sequence of precessional levels showing the characteristic rotational spectrum with energies proportional to $I(I+1)$. In each sequence the lowest level has $I = |K|$ and succeeding levels $I = |K+1|, |K+2|, \ldots$.

Both positive and negative values of the quantum number K are possible, but the rotational energy (4.2) is independent of the sign of K. This degeneracy is removed if the two ends of the rigid body are indistinguishable, i.e. if the body is invariant for rotations through an angle π about the ' 1 ' axis. The rotational wave function should therefore be symmetrized with respect to this rotation giving

$$\phi_{\text{rot}} = \mathfrak{D}_{MK}^{I}(\theta_i)+(-1)^{I-K}\mathfrak{D}_{M,-K}^{I}(\theta_i). \quad \ldots\ldots(4.4)$$

Since for a rigid rotator I is integral the wave function (4.4) is independent of the sign of K and therefore it can be taken as positive.

In the special case of $K = 0$ the symmetrization causes the wave function to vanish if I is odd, hence only even values $I = 0, 2, 4, \ldots$ of the total angular momentum are possible.

4.2. *Dynamics of Nuclear Deformations*

In § 3 we looked at the nucleus from the point of view of the Hartree model and found that away from closed shells the intrinsic particle motion deforms the nucleus, so that it has a non-spherical equilibrium shape. On the other hand

the Hartree wave function has no definite angular momentum, and is unsatisfactory in this sense. Also the Hartree model nucleus has a definite orientation in space, namely along the z axis of a given coordinate system, a situation which clearly does not exist in practice. The solution to this problem lies in treating the orientation of the nucleus as a dynamical quantity. In other words, the nucleus is allowed to rotate. Properly other collective modes should be included in the discussion, principally the possibility of a vibration in the nuclear shape. It turns out, however, that provided the fluctuations in the nuclear deformation are small compared with its equilibrium value then these fluctuations can be neglected in the first instance. In the following sections, therefore, attention will be focused on the nuclear rotations.

As mentioned in the introduction the frequency of rotation of the nucleus is thought to be small compared with the frequency associated with the intrinsic particle motion, so that the rotation of the nucleus should not appreciably disturb the intrinsic motion of the nucleons relative to the deformed shape. In this case the adiabatic approximation is valid and the nuclear wave function can be written as a product of a wave function describing the rotation and a wave function representing the internal nucleon motion. The second part is the wave function for nucleon motion in a deformed nucleus and can be approximated by the Hartree–Fock wave function of § 3. This solution is not logically consistent, however, for originally the problem was specified by $3A$ particle coordinates, while now there are 3 coordinates for the collective rotation and $3A$ more in the intrinsic wave function. The problem of self-consistency has been treated by Villars (1957), Lipkin, de Shalit and Talmi (1955), Peierls and Yoccoz (1957) and others.

The total angular momentum $\mathbf{I}$ of a rotating nucleus is made up of a part $\mathbf{J}$ coming from the intrinsic motion of the nucleons and a part $\mathbf{R}$ coming from the collective rotation. If the nuclear rotations are analogous to that of a symmetric top, the collective Hamiltonian has the form

$$H_{\text{rot}} = \frac{\hbar^2}{2\mathcal{I}_1}(R_1^2+R_2^2)+\frac{\hbar^2}{2\mathcal{I}_3}R_3^2$$

by analogy with (4.1). Here R_1, R_2 and R_3 are the components of $\mathbf{R}$ along the principal axes of the nuclear deformation and $\mathcal{I}_1, \mathcal{I}_3$ are the principal moments of inertia of the deformed nucleus. Making use of the relation $\mathbf{R} = \mathbf{I} - \mathbf{J}$ and expanding (4.5) becomes

$$H_{\text{rot}} = \frac{\hbar^2}{2\Im_1}(I_1^2+I_2^2)-\frac{\hbar^2}{\Im_1}(I_1J_1+I_2J_2)+\frac{\hbar^2}{2\Im_1}(J_1^2+J_2^2)+\frac{\hbar^2}{2\Im_3}(I_3-J_3)^2. \quad \ldots\ldots(4.6)$$

In this expanded expression the second term represents the Coriolis force acting on the nucleons due to the nuclear rotation, often referred to as the rotational particle coupling. The presence of this term implies a breakdown of the adiabatic approximation, thus it is in the spirit of the adiabatic approximation to neglect the rotational particle coupling term in the first instance. The third term in (4.6) affects only the internal particle motion and its effect can be shown to be

small. Leaving out these two terms an approximate form of the rotational Hamiltonian

$$H_{\rm rot} = \frac{\hbar^2}{2\Im_1}(I_1^2+I_2^2)+\frac{\hbar^2}{2\Im_3}(I_3-J_3)^2 \qquad \ldots\ldots(4.7)$$

is obtained, differing only from the symmetric top Hamiltonian (4.1) by a modification of the second term.

As in the discussion of the symmetric top $\mathbf{I}^2$, I_z and I_3 commute with the approximate rotational Hamiltonian (4.7), hence they can be taken as constants of the motion with corresponding eigenvalues I, M and K. In addition the intrinsic particle angular momentum resolved along the nuclear symmetry axis is

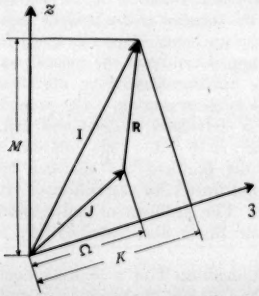


Diagram of the angular momentum coupling scheme in a rotating nucleus. **I**, **R**, and **J** represent the total, collective and particle angular momenta, respectively. K and Ω are the projection of **I** and **J** on the nuclear symmetry axis 3, and M is the projection of **I** along the z axis of a coordinate system fixed in space.

conserved and has eigenvalue Ω. The angular momentum coupling scheme is illustrated by a vector diagram in the figure.

Equation (4.7) indicates that the collective motion of the nucleus will be represented by a wave function which is a rotation matrix $\mathfrak{D}_{MK}^{I}(\theta_i)$ and as mentioned above the intrinsic wave function is assumed to be a Hartree–Fock wave function which we denote by χ_Ω. Because of the symmetry of the nuclear deformation the total wave function must be symmetrized, as discussed in the case of the symmetric top (4.4), and the appropriate form in the adiabatic approximation is

$$\langle r\,|\,\Omega, IMK\rangle = \left[\frac{2I+1}{16\pi^2}\right]^{1/2}[\chi_\Omega \mathfrak{D}_{MK}^{I}(\theta_i)+(-1)^{I-K}\chi_{-\Omega}\mathfrak{D}_{M,-K}^{I}(\theta_i)]. \qquad \ldots\ldots(4.8)$$

The square root factor in (4.8) is a normalization factor arising from the normalization properties of the rotation matrices.†

† The phase of $\chi_{-\Omega}$ is defined by the requirement that $\chi_{-\Omega} = R(0,\pi,0)\,\chi_\Omega$. ($R(0,\pi,0)$ denotes a rotation through an angle π about the '1' axis.) Thus if χ_Ω is the single particle state $\phi_{j\Omega}(\theta,\phi)$, then $\chi_{-\Omega} = (-1)^{j+\Omega}\phi_{j,-\Omega}(\theta,\phi)$.

4.3. *Rotational Energy Levels*

Neglecting the Coriolis term, the rotational energy of a nucleus has the form (4.7)

$$E = \frac{\hbar^2}{2\Im_1}(I^2 - I_3^2) + \frac{\hbar^2}{2\Im_3}(I_3 - J_3)^2$$
$$= \frac{\hbar^2}{2\Im_1}\{I(I+1) - K^2\} + \frac{\hbar^2}{2\Im_3}(K-\Omega)^2. \qquad \ldots\ldots (4.9)$$

All known nuclear spectra are characterized by the requirement $K = \Omega$ or $R_3 = 0$, so that there is no collective rotation about the nuclear symmetry axis. Rotations about the symmetry axis are suppressed either because $\Im_1 \gg \Im_3$ or because of a symmetry effect analogous to that preventing rotations of a spherical nucleus. Thus only the first term of (4.9) contributes to the rotational energy of the nucleus.

All nuclei which are sufficiently deformed for rotation–vibration interaction effects to be small have rotational spectra characterized by the equation (4.9) with $K = \Omega$. Such nuclei are found in the atomic weight regions $A \simeq 24$, $150 < A < 190$ and $A > 222$. The ground state of a nucleus has spin $I_0 = \Omega$, where Ω is the positive value of J_3 of the deformed nucleus shell model state. On top of this there are rotational levels given by (4.9) with $I = \Omega+1, \Omega+2, \ldots$. Corresponding to excited states in the deformed nucleus shell model there are rotational bands in the actual nucleus. Experimental information on these rotational sequences is summarized by Alder *et al.* (1956).

The above remarks break down in two special cases. If $\Omega = I_0 = 0$, as it is in all even–even nuclei, the symmetrization of the wave function (4.8) eliminates the rotational states with I odd. Thus deformed even–even nuclei have an excited state spectra $I = 0, 2, 4, \ldots$ with energies proportional to $I(I+1)$. The ratio of energy of the $I = 4$ second excited state to the $I = 2$ first excited state is $3\frac{1}{3}$. A value close to this is found in most even–even nuclei in the deformed nucleus regions.

If $\Omega = \frac{1}{2}$ the Coriolis term (rotational particle coupling) in (4.6) cannot be neglected. In all other cases the rotational particle coupling has no diagonal matrix elements; but for $I = \frac{1}{2}$ the symmetrization of the wave function (4.8) gives a diagonal contribution to the energy

$$\frac{\hbar^2}{2\Im_1}(-1)^{I+1/2}\langle\chi_{1/2}|J_+|\chi_{-1/2}\rangle\langle I_{1/2}|I_-|I_{-1/2}\rangle = \frac{\hbar^2}{2\Im_1}(-1)^{I+1/2}(I+\tfrac{1}{2})\,a.$$

Here $J_\pm = J_x \pm iJ_y$ and $I_\pm = I_x \pm iI_y$. The matrix element $a = \langle\chi_{1/2}|J_+|\chi_{-1/2}\rangle$ can be calculated directly from the wave functions given by Nilsson (1955) and others and is often termed the decoupling factor.

In the second order the rotation–vibration interaction and the rotation particle coupling give a correction proportional to $[I(I+1)]^2$ to the rotational energies (Kerman 1956).

4.4. *Moments of Inertia*

The moments of inertia associated with the collective rotations of the nucleus can be determined empirically from the observed energy sequences, and it is found that all the observed moments of inertia lie between two limiting cases.

If the nucleus rotates like a rigid body, the moment of inertia about a minor axis is

$$\Im_{\text{rigid}} = \tfrac{2}{5}AMR_0^2(1+\tfrac{1}{3}\epsilon)$$

where R_0 is the radius of an undeformed nucleus of the same volume. On the other hand, if the rotation involves only the nuclear surface, nuclear matter may behave as an irrotational liquid with the deformation as a surface wave running around the drop. In this case the moment of inertia depends more strongly on the deformation with a value

$$\Im_{\text{irr}} = \tfrac{2}{5}AMR_0^2[0{\cdot}89\epsilon^2 + O(\epsilon^3)].$$

Observed moments lie between these limits with values of about $\frac{1}{2}\Im_{\text{rigid}}$ (Bohr and Mottelson 1956).

Theoretical attempts at predictions of moments of inertia have so far been of limited success. Most calculations have been based on the so-called ' cranking model ' of Inglis (1954). The potential well of the deformed nucleus Hartree model is made to rotate with angular velocity W about a minor axis and the motion of the nucleons in this rotating deformed potential is calculated. The rotation carries the nucleons around with it so that the nucleons have a resultant angular momentum L_x about the axis of rotation given in the first order of perturbation by

$$\hbar\langle L_x\rangle_0 = W\hbar^2 \sum_i \frac{|\langle i|L_x|0\rangle|^2}{E_i - E_0}$$

where E_i is the energy of the intermediate state $|i\rangle$. On the other hand, this angular momentum should be given by

$$\hbar\langle L_x\rangle_0 = \Im_1 W$$

where $\Im_1$ is the moment of inertia of the system about the axis of rotation, thus leading to the expression

$$\Im_1 = \hbar^2 \sum_i \frac{|\langle i|L_x|0\rangle|^2}{E_i - E_0}. \qquad \ldots\ldots(4.11)$$

Bohr and Mottelson (1956) and Inglis (1956) have calculated (4.11) for the deformed oscillator model of § 3.1, and find the rigid moment results for the equilibrium deformation. Inglis (1956) has shown that this result must be modified in the case of the generalized oscillator potential (3.5), while Bohr and Mottelson (1956) prove in a special case that resultant particle forces not included in the self-consistent field can reduce the moment of inertia from the rigid value towards the empirical values. Yoccoz (1957) also finds that resultant particle forces tend to reduce the moment of inertia.

4.5. *Electric and Magnetic Moments*

An even–even nucleus in its ground state ($I = 0$) has a constant rotational wave function, hence it is oriented in all directions with equal probability, and the intrinsic quadrupole moment discussed in § 3 averages to zero. For an odd nucleus, however, the intrinsic angular momentum of the particle motion makes up a large part of the ground state angular momentum of the nucleus. Since the deformation is coupled to the intrinsic particle angular momentum it tends to be oriented in the direction of the total angular momentum of the nucleus and precesses

about it, giving a resultant quadrupole moment which is the projection of the intrinsic moment along the total angular momentum vector.

Quantitatively the quadrupole moment is the expectation value of the quadrupole operator

$$Q(2,0) = \sum_1^z (3\cos^2\theta_i - 1)$$

between the nuclear ground state with $M = I$

$$Q_s = \langle \Omega, KII | Q(2,0) | \Omega, KII \rangle. \qquad \ldots\ldots(4.12)$$

This quantity can be evaluated by a method generally useful in the discussion of nuclear moments and transition probabilities in the rotating nucleus model. A tensor operator $\mathbf{T}(\lambda,\nu)$ depending on the particle coordinates and referred to space fixed axes is related to the same operator $\mathbf{T}'(\lambda,\mu)$ referred to the principal axes of the nuclear deformation by a rotation

$$\mathbf{T}(\lambda,\nu) = \sum_\mu \mathfrak{D}_{\nu\mu}^{\lambda}(\theta_i)\,\mathbf{T}'(\lambda,\mu). \qquad \ldots\ldots(4.13)$$

The rotation matrices $\mathfrak{D}_{\nu\mu}^{\lambda}(\theta_i)$ depending on the Euler angles θ_i of the body-fixed coordinates system, are the same as those used to describe the collective rotations in the wave function (4.8). The quadrupole operator in $Q(2,\mu)$ in (4.12) is such a tensor operator referred to space fixed axes, and making use of (4.13) the expression for the quadrupole moment becomes

$$Q_s = \langle \Omega, KII | \sum_\nu \mathfrak{D}_{0\nu}^{2}\, Q'(2,\nu) | \Omega, KII \rangle.$$

Thus the problem is reduced to calculating the quadrupole moment with respect to the axes fixed in the nucleus, and this is just the intrinsic quadrupole moment of § 3.2. If the wave function (4.8) is substituted the cross terms from the symmetrization cancel (except in the case $\Omega = \frac{1}{2}, 1$) and only the terms with $\nu = 0$ contribute, giving

$$\begin{aligned} Q_s &= \frac{2I+1}{8\pi^2}\int (\mathfrak{D}_{IK}^{I})^{+}\,\mathfrak{D}_{00}^{2}\,\mathfrak{D}_{IK}^{I}\, d\Omega^3 \times \langle \chi_\Omega | Q'(2,0) | \chi_\Omega \rangle \qquad \ldots\ldots(4.14)\\ &= \langle I2K0 | I2IK \rangle \langle I2I0 | I2II \rangle\, Q_0. \end{aligned}$$

The Clebsch–Gordan coefficients in (4.14) arise as a special case of a general expression for the integral over Euler angles of a product of three rotation matrices

$$\int (\mathfrak{D}_{M'K'}^{I'})^{+}\,\mathfrak{D}_{\mu\nu}^{\lambda}\,\mathfrak{D}_{MK}^{I}\, d\Omega^3 = \frac{8\pi^2}{2I'+1}\langle I\lambda M\mu | I\lambda I'M' \rangle \langle I\lambda K\nu | I\lambda I'K' \rangle.$$

Explicit expressions for the Clebsch–Gordan coefficients in (4.14)

$$\langle I2K0 | I2IK \rangle = \frac{3K^2 - I(I+1)}{[I(I+1)(2I+3)(2I-1)]^{1/2}}$$

give the well-known expression for the quadrupole moment

$$Q_s = Q_0 \frac{3K^2 - I(I+1)}{(I+1)(2I+3)}$$

for a general state I in the rotational band K or

$$Q_s = Q_0 \frac{I_0(2I_0-1)}{(I_0+1)(2I_0+3)}$$

for the ground state of the band $I_0 = K$.

Magnetic moments of deformed nuclei can be evaluated similarly. The magnetic dipole operator is

$$\mu = \Sigma(g_s \mathbf{s} + g_l \mathbf{l}) + g_R \mathbf{R} \qquad \ldots\ldots(4.15)$$

where the first term is the sum of the intrinsic particle moments, and the last term represents the contribution of the collective rotation of the nuclear charge to the magnetic moment. The gyromagnetic ratio for the collective motion should be approximately $g_R = Z/A$, this being the average charge of a nucleon. Analogous to the case of the quadrupole moment, the projection of the intrinsic moment along the direction of the total angular momentum contributes to the total moment an amount equal to $K\mu_{\text{int}}/(I+1)$, where μ_{int} is the intrinsic moment calculated in § 3.2. The projection factor $K/(I+1)$ can be calculated by the method used for the quadrupole moment (the magnetic dipole operator is a tensor of rank one) or by a more direct method (Nilsson 1955).

The collective contribution can be calculated as follows. In the state $M = I$,

$$R_z = I_z - J_z;$$

thus
$$\langle \Omega, KII | R_z | \Omega, KII \rangle = I - \langle \Omega, KII | J_z | \Omega, KII \rangle. \qquad \ldots\ldots(4.16)$$

The operator J_z is an intrinsic property of the particle motion with the intrinsic value Ω and has the same projection factor as the intrinsic magnetic moment. Thus

$$\begin{aligned} \langle R_z \rangle &= I - K\Omega/(I+1) \\ &= I - K^2/(I+1) \end{aligned}$$

in the usual case of $\Omega = K$. If we write the intrinsic moment in terms of the intrinsic g-factor $\mu_{\text{int}} = g_\Omega K$, the total magnetic moment becomes

$$\mu = \frac{K^2}{I+1}(g_\Omega - g_R) + Ig_R.$$

Thus the collective and intrinsic gyromagnetic ratios can be determined separately by measuring the magnetic moment in the ground and first excited states of a rotational band.

In the case $K = \frac{1}{2}$ the symmetry of the wave function gives rise to an additional term in the magnetic moment in the same way that the decoupling factor of § 4.3 arises in the energy spectrum. This case has been treated by Nilsson (1955).

4.6. *Electromagnetic Transition Probabilities*

In addition to the large quadrupole moments nuclei in the deformed regions have large transition probabilities for emission of electric quadrupole radiation, and again this fact is associated with the large intrinsic quadrupole moments possessed by these nuclei. The transition probability for radiation of a photon of

multipole order λ and frequency $\omega/2\pi$ is given by

$$T(\lambda) = \frac{8\pi(\lambda+1)}{\lambda[(2\lambda+1)!!]^2}\frac{1}{\hbar}\left(\frac{\omega}{c}\right)^{2\lambda+1} B(\lambda)$$

where the reduced transition probability $B(\lambda)$ can be written as

$$B(\lambda) = \sum_{\mu, M_f} |\langle i|\, M(\lambda,\mu)|f\rangle|^2$$

in terms of matrix elements of the multipole operator $\mathfrak{M}(\lambda,\mu)$ of order (λ,μ) between the initial state i and the final state f with magnetic quantum number M_f. The general expressions for $\mathfrak{M}(\lambda,\mu)$ are given, for instance, by Nilsson (1955), but the most important are the magnetic dipole and electric quadrupole operators

$$\mathfrak{M}_m(1,\mu) = \frac{e\hbar}{2mc}\mu(1,\mu)$$

$$\mathfrak{M}_e(2,\nu) = e\left(\frac{5}{4\pi}\right)^{1/2} Q(2,\nu)$$

where $\mu(1,\nu)$ is the tensor form of the dipole operator (4.15) and $Q(2,\nu)$ is the quadrupole operator in (4.12).

4.6.1. *Collective transitions.*

Within a rotational band the matrix elements involved for quadrupole transitions are of the form

$$\langle \Omega, KI_i M_i|\, Q(2\nu)|\, \Omega, KI_f M_f\rangle.$$

This matrix element can be evaluated by the methods of § 4.5 by first making the transformation (4.13) then substituting the collective wave function (4.8), yielding the result

$$Q_0\left[\frac{2I_i+1}{2I_f+1}\right]^{1/2}\langle I_i 2M_i\nu|I_i 2I_f M_f\rangle\langle I_i 2K0|I_i 2I_f K\rangle.$$

Taking the square modulus and summing over final states and directions of emission

$$B_e(2) = \frac{5}{16\pi}Q_0^2\frac{2I_i+1}{2I_f+1}|\langle I_i 2K0|I_i 2I_f K\rangle|^2.$$

The other Clebsch–Gordan coefficients sum to unity by normalization. The reduced transition probability is proportional to the square of the intrinsic quadrupole moment Q_0^2, hence the large value of E2 transition probabilities.

There exist collective M1 transitions involving the operator $\mathfrak{M}_m(1,\nu)$, whose reduced transition probabilities can be calculated by the same method. The results for both the electric quadrupole and magnetic dipole cases are given by Bohr and Mottelson (1953).

4.6.2. *Particle transitions.*

Besides the collective electromagnetic transitions there are particle transitions involving a change of the intrinsic wave function χ_Ω. These have been discussed extensively by Alaga, Alder, Bohr and Mottelson (1955). Most important amongst

this group will be magnetic dipole transitions. In this case the matrix element to be evaluated is of the form

$$\langle \Omega_i K_i I_i M_i | \mathfrak{M}_m(1,\mu) | \Omega_f K_f I_f M_f \rangle = \langle i | \mathfrak{M}_m(1,\mu) | f \rangle .$$

Again making use of the transformation (4.13) this matrix element becomes

$$\langle i | \mathfrak{M}_m(1,\mu) | f \rangle = \langle i | \sum_\nu \mathfrak{D}_{\mu\nu}^{1} \mathfrak{M}_m'(1,\mu) | f \rangle . \qquad \ldots\ldots (4.17)$$

Here $\mathfrak{D}_{\mu\nu}^{1}$ operates only on the collective part of the wave function while $\mathfrak{M}_m'(1,\mu)$ operates only on the intrinsic part. Evaluation of (4.17) yields the result

$$\langle i | M_m(1,\mu) | f \rangle = \frac{2I_f+1}{2I_i+1} \langle I_i 1 M_i \mu | I_i 1 I_f M_f \rangle \langle I_i 1 K_i \nu | I_i 1 I_f K_f \rangle \times \langle \Omega_i | \mathfrak{M}_m'(1,\mu) | \Omega_f \rangle . \qquad \ldots\ldots (4.18)$$

The Clebsch–Gordan coefficients again come from the collective part of the wave function, while the matrix element is between the intrinsic nuclear states of § 3.1, χ_{Ω_i} and χ_{Ω_f}. This will correspond to a one-particle transition if the intrinsic motion is described by deformed shell model wave functions of § 3, and the matrix element can be calculated from the wave functions given by Nilsson (1955), Moszkowski (1955) and Gottfried (1956). It should be noted that the magnetic dipole operator (4.15) can be written as

$$\boldsymbol{\mu} = \Sigma(g_s \mathbf{s} + g_l \mathbf{l}) - g_R \mathbf{J} + g_R \mathbf{I}$$

and that the $-g_R \mathbf{J}$ part of the collective moment is included in the calculation of the intrinsic matrix element. Equation (4.18) gives for the reduced transition probability

$$B_m(1) = \sum_\nu |\langle I_i 1 K_i \nu | I_i 1 I_f K_f \rangle|^2 |\langle \Omega_i | \mathfrak{M}'(1,\nu) | \Omega_f \rangle|^2 . \qquad \ldots\ldots (4.19)$$

Only one term of this sum contributes since $\nu = \Omega_i - \Omega_f$.

From the form of (4.19) it is seen that, in addition to the normal magnetic dipole selection rule $I_f = I_i$ or $I_i \pm 1$, there is a selection rule $\Omega_f = \Omega_i$ or $\Omega_i \pm 1$, and since $\Omega = K$ in all known spectra, $K_f = K_i$ or $K_i \pm 1$. This selection rule is only approximate and holds in the limit that wave function (4.9) is adequate and the rotational particle coupling can be neglected. The breakdown of this selection rule has been discussed by Alaga *et al.* (1955) and by Kerman (1956).

Again, from (4.19), it is clear that reduced transition probabilities from one initial state $I_i K_i$ to two different final states I, I' both in the rotational band K_f have the ratio

$$\frac{B_m(1,I)}{B_m(1,I')} = \left| \frac{\langle I_i 1 K_i \nu | I_i 1 I K_f \rangle}{\langle I_i 1 K_i \nu | I_i 1 I' K_f \rangle} \right|^2 .$$

This relation has been given by Alaga *et al.* (1955) together with examples of its application and experimental verification.

Intrinsic electric quadrupole transitions can be discussed similarly, but now these have transition probabilities characteristic of single particle transitions, much smaller than for the collective transitions.

§ 5. Nuclear Vibrations

Outside the rotational regions and excluding the few nuclei adjacent to closed shells it might be expected that there are regions where the nuclei have a tendency to deform, but where the fluctuations in the nuclear deformation are large compared with its magnitude. If these fluctuations have dynamical properties there should be corresponding excitations analogous to waves on the nuclear surface, a possibility discussed by Bohr (1952).

Collective coordinates $\alpha_{\lambda\mu}$ are introduced by the equation

$$R(\theta,\phi) = R_0\left[1+\sum_{\lambda\mu}\alpha_{\lambda\mu}\,Y_{\lambda\mu}(\theta,\phi)\right]$$

where $R(\theta,\phi)$ is the equation of the deformed nuclear surface. Only deformations with $\lambda \geqslant 2$ are important since $\lambda = 0$ corresponds to a volume vibration which should have a high frequency and $\lambda = 1$ corresponds to a centre of mass translation. In fact, the only possibilities which appear relevant to experiments at the moment are $\lambda = 2$ and $\lambda = 3$.

A Hamiltonian of the form

$$H = \sum_{\mu}(\tfrac{1}{2}B_2|\dot{\alpha}_{2\mu}|^2+\tfrac{1}{2}C_2|\alpha_{2\mu}|^2) \qquad \ldots\ldots(5.1)$$

is chosen to describe $\lambda = 2$ vibrations in spherical nuclei.

Such a Hamiltonian of oscillator form (5.1) leads to a sequence of equally spaced levels with energies multiples of the oscillator energy $\hbar\omega = \hbar(C_2/B_2)^{1/2}$. Deformations of order $\lambda = 2$ have an associated angular momentum $I = 2$. On this model an even–even nucleus has a ground state of spin $I = 0$, a first excited state with spin $I = 2$ and energy $\hbar\omega$, a group of second excited states with spins $I = 0, 2, 4$ degenerate with energy $2\hbar\omega$, and so on. Such a nucleus has a characteristic spectrum for emission of electromagnetic radiation. The ground state and first excited state are connected by an electric quadrupole (E2) transition matrix element enhanced over the single particle value, and similarly between the second excited state group and the first excited state. The crossover E2 transition between the second $I = 2$ state and the ground state is forbidden, however, since it involves a change of two in the oscillator quantum number. The same restriction holds for the crossover M1 transition.

Outside the rotational regions many even–even nuclei with the $I = 0, 2, 4$ sequence of excited states are found with the above characteristic spectra, i.e. equally spaced levels and with the crossover transition strongly inhibited (Scharff-Goldhaber and Weneser 1955).

§ 6. Recent Theoretical Developments

In the above analysis the collective and individual particle aspects of nuclear structure have been considered and brought together into a unified model. The model has been discussed only in the adiabatic approximation, i.e. it is assumed that the rotation does not disturb the intrinsic particle motion. This approximation is valid only so long as the frequency associated with the rotation is small compared with frequencies of particle motions, or equivalently the spacings of rotational levels are small compared with those of particle levels. In practice

this condition is rarely fulfilled and corrections should be made for the effect of the rotation on the intrinsic particle motion. At this point an unsatisfactory aspect of the theory emerges in that collective coordinates have been introduced in § 4.3 in addition to single particle coordinates. The problem is over-specified and there exist subsidiary conditions connecting the collective and particle coordinates which have been ignored in the above analysis. This inconsistency is probably unimportant in the adiabatic approximation, but it becomes a stumbling block in attempts to extend the theory to include rotational particle interactions properly. For instance, at the present stage it is impossible to decide how much of the two-body interaction is already included in the collective analysis and what part should be treated as a perturbation. Also it is probable that the problem of the moments of inertia of deformed nuclei can be solved only by a self-consistent treatment.

Various attempts have been made to put the collective model on a firm theoretical basis. One approach in the direction of self-consistency lies in including only particles outside closed shells in the intrinsic wave function and replaces the particles in closed shells by a rigid core with a certain effective moment of inertia. The model obtained is amenable to exact discussion and is not over-specified, the coordinates of the problem being the coordinates of particles outside closed shells and the orientation of the core. This model has been applied to discussions of rotational particle coupling by Kerman, and the approach has some theoretical justification as a result of investigations of many authors, most recently in work by Villars (1957).

A second and quite different approach has been suggested by Wheeler (1954) and developed by Skyrme (1957), Peierls and Yoccoz (1957), and others. In § 3 a deformed Hartree model wave function was introduced to represent the nuclear wave function, the deformation being chosen to minimize the energy of the model. Obviously the expectation value of the energy for such a wave function is independent of the orientation of its deformation. Let $\chi_\Omega(\theta_i)$ be the Hartree wave function corresponding to a deformation with orientation specified by the Euler angles θ_i, obtained from $\chi_\Omega(0)$ by a rotation, and consider a superposition θ of these wave functions with a weight factor $f(\theta_i)$

$$\phi(r_i) = \int f(\theta_j)\, \chi_\Omega(r_i, \theta_j)\, d\Omega^3. \qquad \ldots\ldots(6.1)$$

Wheeler uses this wave function as a trial function in a variation calculation and $f(\theta_i)$ determined to minimize the energy. On making this variation it is found that

$$f(\theta_i) = D^{I}_{M\Omega}(\theta_i), \qquad \ldots\ldots(6.2)$$

i.e. $f(\theta_i)$ is a rotation matrix element. With this choice $\theta(r_i)$ is an eigenfunction of J^2 and J_z with eigenvalues I and M; thus in the same step the difficulty that the deformed Hartree wave function is not an eigenfunction of total angular momentum has been resolved.

The wave function defined by (6.1) can be used to calculate energies, electric and magnetic moments, transition probabilities, etc. Yoccoz has demonstrated that the model predicts rotation bands with moments of inertia consistent with

experiment. It can be shown further that, provided the intrinsic wave function χ_Ω satisfies certain conditions, this model gives the same predictions for moments and transition probabilities as the simplified adiabatic model discussed in this review. Indeed it seems possible that this variation treatment may give criteria for the validity of the simplified model.

In calculation of energies and matrix elements a quantity of importance appears which is the overlap between two intrinsic wave functions differing slightly in their orientation

$$I(\theta) = \int \chi_\Omega^*(\mathbf{r}_i, \theta)\, \chi(\mathbf{r}_i, 0)\, d\mathbf{r}_i$$

(except for unimportant phase factors $I(\theta)$ depends only on the polar Euler angle θ). Obviously $I(0) = 1$ and calculations made by Yoccoz indicate that $I(\theta)$ is strongly peaked in the neighbourhood of $\theta = 0$. The width of this function determines whether collective motion is strongly developed. One can write down an approximate uncertainty relation for the width $\Delta\theta$ of the overlap function $I(\theta)$ and the range ΔJ of angular momenta included in the intrinsic wave function

$$\Delta J \Delta\theta \sim 1$$

with ΔJ measured in units of $\hbar$. The operation (6.1) picks out a certain component of the intrinsic wave function χ_Ω to represent a rotational state, hence the range of rotational angular momenta is determined by the range ΔJ of angular momenta in the intrinsic wave function. The number of rotational states in a band is of the order of $\Delta J = 1/\Delta\theta$; thus if $I(\theta)$ is narrow rotational modes should be well developed and if the opposite is the case collective rotations cannot exist. Alternatively, $|I(\theta)|^2$ is the probability of finding the intrinsic state oriented at an angle θ with its initial direction—it represents fluctuations in the orientation of a deformed nucleus produced by fluctuations of its structure. Evidently collective rotation can exist only if fluctuations in the intrinsic structure are small.

Acknowledgments

I would like to thank Dr. Blin-Stoyle and Dr. Satchler for helpful suggestions and for reading the manuscript.

References

Alaga, G., Alder, K., Bohr, A., and Mottelson, B., 1955, *K. Danske Vidensk. Selsk. Mat.-fys. Medd.*, **29**, No. 9.

Alder, K., Bohr, A., Huus, T., Mottelson, B., and Winther, A., 1956, *Rev. Mod. Phys.*, **28**, 432.

Bethe, H., 1956, *Phys. Rev.*, **103**, 1353.

Bohr, A., 1952, *K. Danske Vidensk. Selsk. Mat.-fys. Medd.*, **26**, No. 14.

Bohr, A., and Mottelson, B., 1953, *K. Danske Vidensk. Selsk. Mat.-fys. Medd.*, **27**, No. 16; 1954, *Beta and Gamma Ray Spectroscopy*, edited by K. Siegbahn (Amsterdam: North-Holland), chap. 17; 1956, *K. Danske Vidensk. Selsk. Mat.-fys. Medd.*, **30**, No. 15.

Bohr, N., and Kalchar, F., 1937, *K. Danske Vidensk. Selsk. Mat.-fys. Medd.*, **14**, No. 10.

Brueckner, K., 1955, *Phys. Rev.*, **100**, 36.

Elliot, J., 1957, Pittsburgh Conference.

Gallone, S., and Salvetti, C., 1953, *Nuovo Cim.*, **10**, 145.

Goldhaber, M., and Teller, E., 1948, *Phys. Rev.*, **74**, 1046.

GOTTFRIED, K., 1956, *Phys. Rev.*, **103**, 1017.
INGLIS, D. R., 1952, *Phys. Rev.*, **87**, 915; 1954, *Ibid.*, **96**, 1059; 1956, *Ibid.*, **103**, 1786.
JAHN, H. A., and TELLER, E., 1937, *Proc. Roy. Soc.* A, **161**, 220.
KERMAN, A. K., 1956, *K. Danske Vidensk. Selsk. Mat.-fys. Medd.*, **30**, No. 15.
LIPKIN, H. J., DE SHALIT, A., and TALMI, I., 1955, *Nuovo Cim.*, **2**, 773.
MOSZKOWSKI, S., 1955, *Phys. Rev.*, **99**, 803; 1956, *Ibid.*, **103**, 1328.
NILSSON, S. G., 1955, *K. Danske Vidensk. Selsk. Mat.-fys. Medd.*, **29**, No. 16.
PEIERLS, R. E., and YOCCOZ, J., 1957, *Proc. Phys. Soc.* A, **70**, 381.
RAINWATER, J., 1950, *Phys. Rev.*, **79**, 432.
SCHARFF-GOLDHABER, G., and WENESER, J., 1955, *Phys. Rev.*, **98**, 212.
SKYRME, T. H. R., 1957, *Proc. Phys. Soc.* A, **70**, 433.
VAN VLECK, J. H., 1935, *Phys. Rev.*, **48**, 367.
VILLARS, F., 1956, *Nuclear Physics*, **3**, 240.
WEISSKOPF, V. F., 1957, *Rev. Mod. Phys.*, **29**, 174.
WHEELER, J. A., *Proceedings of* 1954 *Glasgow Conference on Nuclear and Meson Physics* (London: Pergamon Press).
YOCCOZ, J., 1957, *Proc. Phys. Soc.* A, **70**, 388.

FOURIER METHODS IN STRUCTURE ANALYSIS BY ELECTRON DIFFRACTION

By J. M. COWLEY and A. L. G. REES

Chemical Physics Section, Division of Industrial Chemistry, Commonwealth Scientific and Industrial Research Organization, Melbourne, Australia

CONTENTS

	PAGE
§ 1. Definitions and nomenclature	165
§ 2. Introduction	167
2.1. Possibility of structure analysis from electron diffraction data	167
2.2. The scattering of electrons by potential fields	171
2.3. Fourier methods of structure analysis	175
§ 3. Intensities in electron diffraction patterns	178
3.1. Atomic scattering factors for electrons	178
3.2. Diffraction by perfect crystals	182
3.3. Diffraction by imperfect crystals	186
§ 4. Structure analysis from intensities	191
4.1. Collection of intensity data	191
4.2. Interpretation of electron diffraction Fourier maps	195
4.3. Structure analysis from polycrystalline material	196
4.4. Single-crystal structure analysis	199
4.5. Applications of structure analysis from intensities	203
§ 5. Structure analysis from dynamic scattering effects	204
5.1. Theory of dynamic effects	204
5.2. Deduction of structure amplitudes from spot-group fine structure	208
5.3. Assessment as a method of structure analysis	213
§ 6. Direct observation of crystal lattice structure	214
6.1. Electron microscope images of lattice structures	214
6.2. Fourier images	216
6.3. Scattering by successive crystals	219
References	224

Abstract. The various methods of deducing the structure of crystals from electron diffraction data are reviewed. A rational and uniform nomenclature is proposed in an effort to resolve the present confusion. A critical review of the theory of electron scattering, including discussion of the nature of the approximations, is given. The application of Fourier methods to the analysis of crystal structure from both diffraction intensities and fine structure of dynamic origin is outlined and illustrated for polycrystalline and single-crystal specimens. Some remarks are made on the possibilities of direct observation of crystal structure by electron-optical methods.

§ 1. Definitions and Nomenclature

Nomenclature in the field of electron diffraction is, at the moment, in a rather unsatisfactory state. Different authors use the same symbol for entirely different quantities and often the same names and symbols are applied to quantities differing by a multiplicative constant, not always dimensionless. In an effort to resolve the confusion over symbols and nomenclature, we propose here a system of nomenclature which retains as much of existing usage as is consistent with the following set of rules.

(i) The accepted symbols used in x-ray diffraction work should be retained when the quantities described are identical.

(ii) For quantities in electron diffraction analogous to, but not identical with, quantities in x-ray diffraction, the symbols should not be the same, but some relationship (e.g. alphabetic proximity) is desirable. The distinction should not be made by means of subscripts or superscripts.

(iii) Functions of real space coordinates are to be represented by Greek letters. Functions of reciprocal lattice space are to be represented by italic capitals if referring to crystals or by small italic letters if referring to individual atoms.

(iv) The relevant quantities should be defined in such a way as to reduce the number of multiplicative constants in the more common equations to a minimum.

The symbols used in this article and proposed for general use are:

x, y, z real space coordinates. u, v, w reciprocal space coordinates. h, k, l reciprocal lattice indices. a, b, c unit cell dimensions. V unit cell volume. W accelerating potential of electron beam. $\phi(x, y, z)$ potential distribution in real space. $E(u, v, w)$ the Fourier transform of $\phi(x, y, z)$. $E(h, k, l)$ or E_{hkl} the Fourier coefficients for periodic $\phi(x, y, z)$. $f(\theta)$ atomic scattering factor for x-rays. $g(\theta)$ atomic scattering factor for electrons, given in the Born approximation by the Fourier transform of $\phi(x, y, z)$ for a single atom.†

The equations of particular interest for purposes of structure analysis are then

$$\phi(xyz) = \frac{1}{V}\sum_h \sum_k \sum_l E_{hkl} \exp\{-2\pi i(hx+ky+lz)\}$$

$$E_{hkl} = \sum_j g_j \exp\{-2\pi i(hx_j+ky_j+lz_j)\},$$

which are completely analogous to the corresponding equations for the x-ray case,

$$\rho(xyz) = \frac{1}{V}\sum_h \sum_k \sum_l F_{hkl} \exp\{-2\pi i(hx+ky+lz)\}$$

$$F_{hkl} = \sum_j f_j \exp\{-2\pi i(hx_j+ky_j+lz_j)\}.$$

In the first Born approximation the atomic scattering factor $g(\theta)$ is given by

$$g(\theta) = \frac{e}{4\pi}\{Z - f(\theta)\} \Big/ \left(\frac{\sin^2\theta}{\lambda^2}\right)$$

$$= f_{el}(\theta)\, h^2/2\pi me,$$

where f_{el} (in cm) is the atomic scattering factor as defined by Pinsker (1949) and Weinstein (1953), who tabulated

$$f_{el}(\theta)(4\pi^2 h^2/8\pi^2 me^2) \times 10^{-2} = g(\theta)(\pi/e) \times 10^{-2},$$

and also by Ibers and Hoerni (1954), who calculated complex atomic scattering factors $f_{el}(\theta)$, and tabulated $f_{el}(\theta) \times 10^8$, i.e. the values in ångström units.

The structure factor E_{hkl} is identical with the Φ_{hkl} used by Pinsker and his colleagues and differs only by a factor of the unit cell volume V from the V_{hkl} used by Bethe (1928), MacGillavry (1940) and others.

† To simplify the mathematical expressions we have used $\phi(xyz)$, etc., for $\phi(x, y, z)$, etc. in places where no ambiguity can arise.

§ 2. Introduction

2.1. *Possibility of Structure Analysis from Electron Diffraction Data*

That intensities of electron diffraction patterns could serve as a basis for structure analysis of small crystals has probably been appreciated since the principal theoretical predictions regarding the wave properties of electrons were confirmed experimentally. The scattering of electrons by crystals was shown to be qualitatively similar to the scattering of x-rays, which itself had been established as a basis for structure analysis. However, early attempts to use electron diffraction intensities for this purpose were restricted and incomplete (see Pinsker 1953 for references to this early work), and the first full-scale use of electron diffraction data for structure analysis by Fourier series techniques was made by Pinsker, Weinstein and their colleagues at the Institute of Crystallography of the Academy of Sciences of the U.S.S.R., Moscow, some twenty years after the classical experiments of Davisson, Germer, Thomson and Reid. The Russian work was, at least in the beginning, confined to the use of polycrystalline specimens giving ring patterns or patterns of arcs arranged on a series of near parallel layer lines. The use of single-crystal data for the purpose of structure analysis was first explored and subsequently developed by a group in the Chemical Physics Section, C.S.I.R.O. Division of Industrial Chemistry, Melbourne.

Ignorance of the factors determining the distribution of intensity in electron diffraction patterns is probably responsible for the long delayed exploitation of electron diffraction for structure analysis. Many anomalous intensity distributions were recorded; simple metals gave inconsistent intensities in ring patterns; other materials, such as zinc oxide smoke particles, gave results which were self-consistent but different from those calculated from the known structure.

Some of these anomalies were explained in terms of dynamic scattering of electrons in perfect crystals, following Bethe's (1928) theory of electron diffraction. Thus, Blackman (1939) deduced that dynamic, or multiple coherent, scattering of electrons in a perfect crystal should become important for quite small crystal thicknesses. For the strongest reflection of a particular crystal he calculated that a 5% departure from the single-scattering (kinematic) intensity would occur for thicknesses of 46 Å for Al, 18 Å for Ag and 11·5 Å for Au for 37·5 kv electrons and that marked deviations would occur for 180 Å of Al, 70 Å of Ag and 45 Å of Au. Some intensity measurements on ring patterns from metal films appeared to confirm these predictions. However, early intensity measurements could not be made with precision and crystal thicknesses were commonly taken as the average calculated from the mass of metal deposited per unit area. Consequently, the impression that electron diffraction intensities were normally seriously modified by dynamic scattering became widespread, and few attempts to draw positive conclusions from intensity data were made.

The developments in structure analysis by the Moscow group were accompanied by a careful theoretical and experimental re-assessment of the range of crystal thickness for which the kinematic theory can be assumed to be valid. Weinstein (1955 a) has pointed out that crystals of interest for purposes of structure analysis frequently have more atoms per unit cell, lower symmetry and lower average atomic number than the face-centred cubic metals with which

Blackman dealt. Each of these factors tends to increase the critical thickness for which the kinematic theory ceases to be valid. Weinstein calculated that, while this critical thickness is of the order of 50 Å for average reflections of Au and 200 Å for Al it may be as high as 1200 Å for complex structures with 20 to 25 atoms per unit cell and low atomic numbers, 500 to 600 Å for similar structures containing atoms of medium atomic number, and 250 to 400 Å if heavy atoms are present. Such estimates of thickness must be referred to the size of the regions of perfect lattice structure. Since in small crystals the density of imperfections in crystal lattices appears to be at least as great, and frequently much greater than for crystals which grow to appreciable size, the size of perfect crystal regions must often be considerably smaller than the apparent crystal size.

In a series of measurements of intensities of ring patterns given by evaporated metal films, Yamzin and Pinsker (1949) made careful experimental checks on such theoretical predictions. The crystal dimensions in directions perpendicular to the beam, as estimated from ring broadening, were used as an indication of crystal thickness in preference to measurements of average film thickness. It appeared that the intensities agreed with the predictions of the kinematic theory within experimental error for films with crystal dimensions of several hundred Å, except that some anomalies, consistent with a dynamic scattering interpretation, could be observed for the strongest reflections of Au films. Very careful measurements of ring pattern intensities by Lennander (1954) using a Geiger counter technique showed somewhat larger deviations from the predictions of the kinematic theory for comparable metal films. Kuwabara (1957) has made careful photographic measurements which are in agreement with Lennander's results and Blackman's predictions. The evidence is thus conflicting, but it seems to indicate that Blackman's estimates are of the right order. Recently two methods have been proposed whereby the presence of some dynamic scattering effects in electron diffraction patterns can be detected and a correction to the intensities can be made. One method proposed by Honjo and Kitamura (Miyake 1957) involves determining the dependence of the relative intensities on the wavelength of the electrons. The other method described by Nagakura (1957), and also mentioned by Weinstein (1956 c) makes use of the fact that the statistical distribution of intensities with angle, as plotted by Wilson (1949) for the x-ray case, will be changed in a predictable way when some dynamic scattering is present.

With the development of electron microscopy and of high resolution electron diffraction techniques, the additional information which could be obtained on the nature of the diffracting crystals allowed a number of outstanding intensity anomalies to be satisfactorily explained. Zinc oxide smoke particles were shown to be groups of long needle-like crystals. The form of cupric oxide which gives a pattern so different from that of normal CuO that it had been referred to as a new crystal form, CuO′, likewise was shown to be acicular. Cowley (1954) demonstrated that the anomalous intensity distributions in patterns of CuO and ZnO could be adequately accounted for in terms of the acicular habit, the presence of one or more screw dislocations parallel to the long dimension of the crystals, and a range of crystal thickness covering the region in which dynamic scattering effects became appreciable.

There remain very few anomalies in the intensity distributions of ring or arc patterns from polycrystalline materials which have not received an adequate explanation. Criteria in quantitative form are available to determine whether the intensities might be modified by such factors as crystal size, morphology or lattice defects. Provided that such criteria are applied with sufficient care, there can be very little objection to the use of the intensities from such material for purposes of structure analysis.

As in x-ray diffraction work, the use of single-crystal patterns offers considerable advantages over the use of patterns from polycrystalline materials, but the requirements in the matter of experimental technique are more exacting. The intensities of single-crystal spot patterns have commonly been considered even less useful than those of ring patterns. Patterns often showed an almost uniform distribution of intensities among the spots except for the separation of the spots at high diffraction angles into circular Laue zones. There appeared to be not even a qualitative relationship between the intensities observed and those predicted by the kinematic theory. ' Forbidden ' spots, for which space-group considerations require that the kinematic intensity should be zero, appeared no less strongly than other spots.

Such patterns, obtained by reflection methods from large single crystals, were mostly produced by crystals of thickness ranging from zero to the maximum depth of penetration of electrons in the crystal, determined by the inelastic scattering cross section. This maximum thickness is many times greater than the critical thickness for which dynamic scattering becomes appreciable. Normally, also, the projections from the crystal giving rise to the pattern were imperfect, containing lattice faults which, as will be shown later, may well have seriously modified the intensity distributions. The patterns obtained therefore represented the summation of the intensity distributions of a large number of crystal projections, each extensively modified by the effects of dynamic scattering, crystal imperfections and secondary elastic scattering.

The single-crystal patterns obtained by transmission methods were commonly those given by crystals of such materials as mica, for which it was possible to obtain very thin flakes sufficiently extensive to intercept an appreciable fraction of the electron beams used. The electron beams usually had diameters of 0·1 mm or more, particularly when a single lens was used to focus the electron source on the photographic plate. The crystal flakes were frequently distorted and disordered, so that the relation between the observed and theoretical intensities was obscured.

It was not until high resolution techniques were used to obtain electron beams with diameters of only a few microns at the specimen level that the probability of getting patterns from small, undistorted crystals thin enough to avoid dynamic scattering effects became appreciable. Electron microscope observations have indicated that crystals of flaky habit, such as mica and molybdic oxide, thin enough to approximate to kinematic scattering, are almost invariably bent or otherwise distorted if more than a micron or two in diameter. It is therefore necessary to restrict the beam diameter to something of this order to obtain patterns of reasonable intensity.

Such small beam diameters are readily realized by using a lens to focus an electron beam from a small source on to the specimen. In so doing the resolution

of the diffraction pattern is adversely affected. This, however, is not necessarily a disadvantage if the beam divergence is kept so low that the diameters of the diffraction spots are very much less than the separations of the spots, e.g. if the angle of divergence is less than about 10^{-3} radian for relatively small unit cell dimensions.

The introduction of the three-stage electron microscope with two projector lenses has provided an additional method for examining small crystals. In this instrument a portion of the electron microscope image a few microns in diameter may be selected by means of an aperture in the image plane of the objective lens. The diffraction pattern of the selected area may then be obtained by using the first projector lens to image the back focal plane of the objective lens on to the object plane of the second projector lens which then gives a magnified image of it on the screen. A great advantage of this system is that it allows the diffraction pattern to be correlated with the microscope image. The disadvantages include the limitations on specimen manipulation and treatment imposed by the requirement for electron microscopy that the specimen be very close to or within the central part of the field of the objective lens, and the possible distortion of the diffraction pattern in the process of imaging it by means of the projector lenses.

For all single-crystal structure analyses carried out to date, the crystal has been held stationary relative to the incident electron beam. The intensities then recorded correspond to sections by the Ewald sphere of the regions of scattering power about the reciprocal lattice points. In this they differ from the intensities of reflections in patterns from polycrystalline material which correspond to integrations over the regions of scattering power. Moving-beam, or moving-crystal, techniques to give integrated intensities in single-crystal patterns have not yet been applied.

The use of stationary-crystal intensities presupposes that there is a monotonic decrease of scattering power around the reciprocal lattice points. It also requires that the orientation of the crystal be adjusted so that the beam is very closely parallel to a principal axis in order that the intensities of the spots should be symmetrical about the central spot. This imposes considerable experimental difficulties.

The single-crystal methods, and, to a lesser extent, the polycrystalline methods, still need considerable development before they can match the x-ray diffraction techniques in accuracy and reliability. However, with each electron diffraction structure analysis comes improvement in technique and advance in the understanding of the relevant factors.

A comprehensive account of the procedures and results of structure analysis by electron diffraction methods, particularly the polycrystalline methods, is given in a recently published book by Weinstein (1956 c).

Intensity data are not the only possible basis for the structure analysis of crystals by the use of electrons; other approaches exist for which no useful analogue seems to exist for x-rays. For example, structure amplitudes may be deduced from the geometry of the diffraction features produced by crystals of regular habit thick enough for dynamic scattering effects to be appreciable. The possibility of using the information gained from such linear measurements for purposes of structure analysis are discussed in § 5.

Direct imaging of the atoms in a crystal by electron microscopy has not yet been realized. Menter's (1956 a) observations of the grosser features of the unit

cell contents in the case of some metal phthalocyanine crystals have been encouraging in this direction. Theoretical estimates indicate that individual atoms could be seen in an electron microscope if a resolution of 2–3 Å could be attained. Even if the serious experimental difficulties in attaining such resolution are overcome, it is possible that problems may be met in the interpretation of the images. In direct electron microscopy, it would be necessary to use an instrument capable of resolving individual atoms in order to see the relative atomic positions in the periodic array of a crystal lattice. The prior knowledge of the periodicity of a crystal may possibly be used, either to increase the resolution attainable with a given experimental effort, or else to obtain an equivalent resolution with simpler apparatus. Several schemes whereby this might be accomplished are discussed in § 6.

2.2. *The Scattering of Electrons by Potential Fields*

When a beam of electrons enters the potential field associated with the charged particles contained in atoms, molecules or crystals, there are two principal effects. In terms of the wave description of electrons, these are a change of wavelength of the electron beam, associated with a change of refractive index, and reduction in amplitude of the wave, with a corresponding reradiation of waves of longer wavelength, due to interaction with the particles present in the field. If the electrons are regarded as particles the effects are an acceleration of the electrons by the potential field, resulting in elastic collisions, and inelastic collisions of the incident electrons with the particles of the substance. For our present purposes we make use of the wave description and consider the scattering of electrons as a problem in physical optics. We apply some of the basic ideas of the new formulation of physical optics recently described by Cowley and Moodie (1958), and its application to electron diffraction problems (Cowley and Moodie 1957 d).

As is usual in the treatment of electron and x-ray diffraction problems, we assume that the incident radiation is perfectly coherent, both in a geometric and a chromatic sense. Geometrically, we consider that the radiation comes from an ideal point source, or else is parallel, with cophasal planes perpendicular to the direction of propagation. The way in which the intensity distribution of a diffraction pattern is modified by the use of an ideally incoherent source of finite size, which may be considered a reasonable approximation to a practical electron source, is described in another publication (Cowley and Moodie, to be published). It is shown in a paper by Gabor (1956) that the assumption of chromatic coherence is justifiable for most experimental situations. The finite spread of electron wavelengths and the small energy losses involved in elastic scattering processes can be ignored. On the other hand, electrons which have lost an appreciable amount of energy by inelastic scattering may be considered incoherent with the incident radiation, and may be ignored except in that they provide a diffuse, monotonically decreasing background to diffraction patterns. The decrease in the intensity of the chromatically coherent wave, due to inelastic scattering, does modify the coherent scattering distribution, but Boersch (1948) has shown that the contribution to the diffraction pattern due to this cause is negligible compared with the contribution due to the modification of the wave front by the potential field. We therefore ignore this ' absorption ' term.

We consider the coherent incident electron beam travelling in the direction of the z-axis through a potential field $\phi(xyz)$. The wavelength is given by

$$\lambda = h/[2m\{W+\phi(xyz)\}]^{1/2} = \lambda_0\{1+\phi(xyz)/W\}^{-1/2},$$

where λ_0 is the wavelength for $\phi(xyz) = 0$.

We now assume that in a thin slice of the potential field of thickness Δz, centred about $z = z_1$, the potential function does not vary appreciably with z and may be written $\phi(xyz_1)$. Electron waves passing through this slice will undergo a change of phase, relative to the wave in the absence of a field, of amount

$$[\{1+\phi(xyz_1)/W\}^{1/2}-1]\,2\pi\Delta z/\lambda.$$

For $\phi(xyz_1) \ll W$, this is approximately $\pi\Delta z\phi(xyz_1)/\lambda W$. The effect of this slice of the potential field on the wave function is thus given by multiplying the wave function by

$$q(xyz_1) = \exp[i\sigma\phi(xyz_1)\Delta z],$$

where

$$\sigma = \pi/\lambda W = 2\pi me\lambda/h^2. \qquad \ldots\ldots(2.1)$$

The effect of the whole potential field on the wave function can be approximated to by considering the effect of a large number of such slices, together with the effect of propagation of the wave between the centres of the slices. This has been done in detail in the development of the theory of the diffraction of electrons by Cowley and Moodie (1957 d). Propagation of the wave function from one plane to another involves Fresnel diffraction effects. If, however, the total thickness of the potential field in the beam direction is small and the wavelength is very small, the Fresnel diffraction effects can be ignored as a first approximation. Then the total effect of the potential field on the wave function is given by multiplying by

$$\prod_n q(xyz_n) = \exp\left[i\sigma\sum_n \phi(xyz_n)\Delta z\right]$$

or, in the limit for the number of slices tending to infinity, and the thickness of each to zero,

$$q(xy) = \exp[i\sigma\phi(xy)], \qquad \ldots\ldots(2.2)$$

where

$$\phi(xy) = \int\phi(xyz)\,dz.$$

For plane parallel incident radiation $q(xy)$ will then represent the wave function on the exit surface of the field. A diffraction pattern can be defined in terms of the angular distribution of amplitude on a plane at infinity, or, more conveniently, the positional distribution of amplitude on the back focal plane of an ideal thin lens. In either case it can be shown (Cowley and Moodie 1958) that the diffraction pattern amplitude distribution is given by the Fourier transform of the amplitude distribution on the exit face of the specimen. This is, in our approximation, the Fourier transform $Q(uv)$ of $q(xy)$ defined by

$$Q(uv) = \iint q(xy)\exp\{-i(ux+vy)\}\,dx\,dy.$$

If now we assume that $\sigma\phi(xy) \ll 1$, the exponential of (2.2) may be expanded as

$$q(xy) \sim 1 + i\sigma\phi(xy). \qquad \ldots\ldots(2.3)$$

Then $$Q(uv) = \delta(uv) + i\sigma E(uv), \qquad \ldots\ldots(2.4)$$

where $E(uv) = E(uv0)$,

$E(uvw)$ is the Fourier transform of $\phi(xyz)$,

and $\delta(uv)$ a Dirac delta function, representing a sharp peak at the origin.

The intensity distribution in the diffraction pattern is then given by

$$I(uv) = |Q(uv)|^2 = \delta(uv) + |\sigma E(uv)|^2. \qquad \ldots\ldots(2.5)$$

The delta function represents the transmitted, undiffracted beam. The other term represents the intensity distribution of the diffraction pattern as given by the usual, kinematic, theory of electron diffraction for very small wavelengths.

The function $E(uvw)$ may be regarded as a function of reciprocal or Fourier transform space. Then the amplitude distribution in the diffraction pattern $E(uv)$ may be regarded as the section of the reciprocal space function by the plane $w = 0$.

In the normal development of kinematic diffraction theory, or in the theory of Cowley and Moodie (1957 d) for a higher approximation than that which we have been considering, the plane $w = 0$ is replaced by the surface of a sphere of radius $1/\lambda$, passing through the origin and tangential to $w = 0$; the so-called Ewald sphere. The distribution of amplitudes in the diffraction pattern may then be considered as given by the intersection of the Ewald sphere with the function $E(uvw)$.

A rather more common concept is to consider the function $|E(uvw)|^2$, known as the scattering power, as a function of reciprocal space coordinates. The intersection of the Ewald sphere with this function then gives a quantity proportional to the intensity distribution in the diffraction pattern, which is the observable quantity. By making intensity measurements on diffraction patterns obtained for suitably chosen orientations of the Ewald sphere, i.e. for suitably chosen directions of the incident beam relative to the reciprocal space axes, it is possible to deduce experimentally the distribution of scattering power in reciprocal lattice space, $|E(uvw)|^2$. In most applications of diffraction processes, the crucial problem is to deduce the real-space distribution $\phi(xyz)$ from this experimentally determined distribution of scattering power. This is, in general, not possible.

The problem becomes much more tractable, although still not generally soluble, when it is possible to make the assumption that $\phi(xyz)$ is perfectly periodic, or else can be represented by a periodic function multiplied by a step function which falls sharply from unity to zero at the boundaries of a region large in comparison with the periodicities. This sort of approximation is reasonable for many cases of the diffraction of electrons by crystals.

If $\phi_p(xyz)$ is considered perfectly periodic, the functions $E_p(uvw)$ and $|E_p(uvw)|^2$ are sets of delta functions. Thus

$$E_p(uvw) = \sum_h \sum_k \sum_l E_{hkl}\, \delta(u-h, v-k, w-l), \qquad \ldots\ldots(2.6)$$

so that $$\phi_p(xyz) = \frac{1}{V} \sum_h \sum_k \sum_l E_{hkl} \exp\{-2\pi i(hx+ky+lz)\}. \qquad \ldots\ldots(2.7)$$

The set of delta functions defined by the integral coordinates *hkl* form the reciprocal lattice of the crystal. The reciprocal lattice may be defined by the reciprocal vectors $\mathbf{a}^*, \mathbf{b}^*, \mathbf{c}^*$, which are related to the unit cell vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$, of the crystal lattice by the equations

$$\left.\begin{aligned}\mathbf{a}^* &= \mathbf{b}\times\mathbf{c}/V,\\ \mathbf{b}^* &= \mathbf{c}\times\mathbf{a}/V,\\ \mathbf{c}^* &= \mathbf{a}\times\mathbf{b}/V.\end{aligned}\right\} \qquad \ldots\ldots(2.8)$$

For crystals of finite size, $\phi(xyz)$ may be represented by a perfectly periodic function, $\phi_p(xyz)$, multiplied by a shape function, $\tau(xyz)$, which is unity within the crystal boundaries and zero outside. Then the amplitude distribution in reciprocal space is

$$E(uvw) = E_p(uvw) \star S(uvw), \qquad \ldots\ldots(2.9)$$

where $S(uvw)$ is the 'shape transform', or the Fourier transform of the shape function $\tau(xyz)$. The $\star$ symbol represents a convolution, defined by

$$f(x)\star g(x) = \int_{-\infty}^{+\infty} f(X)\,g(x-X)\,dX, \qquad \ldots\ldots(2.10)$$

which implies that the function $g(x)$ is laid down at all points, having a weight $f(X)$ when its origin is displaced by an amount X, and then the whole is summed.

In the present case, since $E_p(uvw)$ is a set of delta functions, $E(uvw)$ is given by replacing each delta function by the function $S(uvw)$, given the relative weight E_{hkl}. If the shape transform functions centred on adjacent reciprocal lattice points do not overlap appreciably, the distribution of scattering power in reciprocal space is given by

$$|E(uvw)|^2 = |E_p(uvw)|^2 \star |S(uvw)|^2. \qquad \ldots\ldots(2.11)$$

For crystals of reasonably large size in terms of unit cell dimensions, $|S(uvw)|^2$ will have a maximum at the reciprocal lattice point position of half-width very much smaller than the distance between reciprocal lattice points, and will fall off rapidly in value, possibly through a series of maxima and minima. The intensity distribution in the diffraction pattern is then given by the intersection of the Ewald sphere with the regions of scattering power about the reciprocal lattice points. Usually in x-ray diffraction from single crystals, the crystal size is so large, and the scattering power regions so small, that in general the Ewald sphere will pass through only one region of appreciable scattering power at a time. In electron diffraction, on the other hand, at least one dimension of the crystal is usually very small, so that the region of scattering power about the reciprocal lattice points is considerably extended in at least one direction. The Ewald sphere may then intersect a large number of such regions when the incident beam is suitably oriented, giving rise to extensive patterns of spots. Because of the small wavelength of the electrons normally used in diffraction experiments, the Ewald sphere radius $(1/\lambda)$ is large and, as a first rough approximation, the Ewald sphere can be assumed planar over a limited region of reciprocal space. If, for the appropriate orientation of the incident beam, the Ewald sphere is approximated by the plane $w = 0$, this plane passes through all reciprocal lattice points $hk0$. The diffraction pattern is then a pattern of spots, each spot representing a central section of the region of scattering power about the reciprocal

lattice points. The intensity of the $hk0$ spot, integrated over the plane of observation, is then proportional to $|E_{hk0}|^2$.

Usually it is necessary to take the finite radius of the Ewald sphere into account. For the same incident beam orientation, the Ewald sphere then is seen to cut the regions of scattering power further and further from the $hk0$ points as the distance from the reciprocal lattice origin is increased. The effect of the consequent reduction of the recorded intensities of outer spots of the pattern has been discussed by Cowley (1953 b).

If an electron diffraction instrument of high resolving power is used, it is sometimes possible to record the detailed intensity distribution of a diffraction spot, and so obtain information concerning the shape transform function $|S(uvw)|^2$. For example, Rees and Spink (1950 a) recorded fine structure, consisting of a line of subsidiary maxima about a central maximum, in diffraction spots given by single crystals of ZnO smoke (figure 21). The detailed intensity distribution was shown to be consistent with the acicular shape and the dimensions of the crystals as observed in the electron microscope.

2.3. *Fourier Methods of Structure Analysis*

From experimental observations of the intensities of spots, arcs or rings in electron diffraction patterns, it is possible to deduce the distribution of scattering power in the reciprocal lattice space of a crystal. We defer discussion of the accuracy and limitations of the experimental techniques and, assuming that all the required information on the scattering power distribution is available, consider how this information may be used to derive the structures of the crystals concerned. The problem is almost identical with that of the structure analysis of crystals by x-ray diffraction methods. Although the experimental techniques of x-ray diffraction are very different, these techniques can also, ideally, lead to a description of the distribution of scattering power in reciprocal lattice space. With only a few exceptions, the whole range of interpretative and manipulative techniques developed for the Fourier method of structure analysis from x-ray diffraction data can be applied to structure analysis from electron diffraction data.

If we do not concern ourselves with any question of the shape of the crystal or the nature or distribution of defects of the crystal lattice, we need to know only one numerical value per reciprocal lattice point, namely the value of the scattering power integrated over the vicinity of the reciprocal lattice point. The complete set of such values contains all the information concerning the average unit cell content available from diffraction experiments for a given crystal and a given incident radiation.

By recording the intensities of diffraction spots, the relative scattering amplitudes of the various diffracted beams can be found, but no information can be gained concerning the relative phases of the diffracted beams. If the relative phases were known, the crystal structure could be deduced immediately by summation of the Fourier series (2.7),

$$\phi(xyz) = \frac{1}{V}\sum_h \sum_k \sum_l E_{hkl} \exp\{-2\pi i(hx+ky+lz)\},$$

where the structure factors E_{hkl} are in general complex. This can be written

$$\phi(xyz) = \frac{1}{V}\left[E(000) + 2\sum_{h=0}^{+\infty}\sum_{k=-\infty}^{+\infty}\sum_{l=-\infty}^{+\infty}|E_{hkl}|\cos\{2\pi(hx+ky+lz)-\alpha(hkl)\}\right], \quad \ldots\ldots(2.12)$$

where $|E_{hkl}|$ are the structure amplitudes which may be deduced from the intensities, and the $\alpha(hkl)$ are the phase angles.

From the knowledge of the structure factors for points in only one plane of reciprocal lattice space it is possible to deduce the corresponding two-dimensional projection of the structure

$$\phi(xy) = \int\phi(xyz)\,dz = \frac{1}{A_c}\sum_h\sum_k E_{hk0}\exp\{-2\pi i(hx+ky)\}, \quad \ldots\ldots(2.13)$$

where A_c is the area of the projection of the unit cell.

Since the relative phases of the diffracted beams cannot be deduced by norma diffraction procedures, other ways of obtaining this, or equivalent, information must be found. A comprehensive account of the methods used to attack this ' phase problem ' in x-ray diffraction work is given in the book by Lipson and Cochran (1953). Most of these methods may be applied without any fundamental modification to electron diffraction structure analysis. We shall give only a brief outline of them here.

The quantities which can be deduced directly from the measured intensities are the squares of the structure amplitudes, $|E_{hkl}|^2$ or $E_{hkl}E_{hkl}^*$. If a Fourier series similar to (2.7) is summed using these quantities as coefficients we obtain the so-called Patterson function,

$$P(XYZ) = \frac{1}{V}\sum_h\sum_k\sum_l E_{hkl}E_{hkl}^*\cos\{2\pi(hx+ky+lz)\}. \quad \ldots\ldots(2.14)$$

By the well-known theorem, used above in (2.9), the Fourier transform of a product of two functions is the convolution of the Fourier transforms of the two functions. Applying this theorem to the case of periodic functions (2.14) then becomes

$$P(XYZ) = \phi(XYZ)\star\phi(-X,-Y,-Z). \quad \ldots\ldots(2.15)$$

Thus, if we consider the potential distribution as made up of well-defined atomic peaks, the Patterson function is made up of the centro-inverted structure laid down with the origin point placed in turn at each atomic position of the structure weighted according to the relative strength of the corresponding atomic peak. If the atomic positions in the unit cell may be represented by the vectors $\mathbf{r}_n$, the atomic positions for the centro-inverted structure will be $-\mathbf{r}_m$, and the Patterson function will have maxima at positions given by $\mathbf{r}_n-\mathbf{r}_m$ of strength $\phi_n\cdot\phi_m$ where ϕ_n and ϕ_m are the strengths of the corresponding potential maxima in $\phi(xyz)$. The Patterson function therefore can be regarded as an indication of the interatomic vectors present in the structure and the relative weights and frequencies of their occurrence. This interpretation has been the basis of its use in various trial-and-error methods for locating prominent features of the atomic arrangement.

A more complete use of the information contained in a Patterson function is possible if the techniques of ' image seeking ' are applied. For example, a polygon

whose corners represent atomic positions in the structure is superimposed on the Patterson function, and the values of the Patterson function at the polygon vertices are added together, multiplied together, or else, preferably, the minimum of the values is chosen. If the result is plotted as a function of the position of a given point of the polygon relative to the Patterson function origin, a function will be generated which resembles the actual structure more closely the greater the number of correctly chosen polygon points. Image-seeking procedures are usually carried out on two-dimensional projections of the Patterson function. By analogy with (2.13) we have

$$P(XY) = \frac{1}{A_c} \sum_h \sum_k |E_{hk0}|^2 \cos\{2\pi(hX + kY)\}. \qquad \ldots\ldots(2.16)$$

In x-ray diffraction work a large number of structures have been analysed by comparing the diffraction patterns of two crystals which differ in structure only by the isomorphous replacement of one or more atoms in the unit cell by atoms of considerably different atomic numbers. The same methods could be applied equally well in electron diffraction work.

No direct analogy exists for electrons to the x-ray techniques making use of the change of phase of scattering by one or more atoms when the incident radiation is so chosen that the wavelength is close to the absorption edge for the element concerned. However, for crystal thicknesses for which the kinematic theory of diffraction can no longer be applied, there is a phase difference between the contributions of light and heavy atoms, dependent on the crystal thickness and the accelerating voltage, which could possibly be made use of in a similar way (see § 3.1).

The direct methods of determining the phases of structure factors by the use of inequality relations between the structure factors, or of probable equality relationships between either the structure factors or their signs, could also be used in the case of electron diffraction, although a certain amount of caution is necessary because of uncertainties in the atomic scattering factor for electrons.

The final criterion for the correctness of a structure analysis is that intensities or scattering amplitudes calculated for the proposed structure should agree with those found experimentally. Theoretical structure factors are usually calculated by making the assumption that the potential distribution in the unit cell of a crystal is given by summing the potential distributions which would be associated with the individual atoms considered separately, i.e.

$$\phi(xyz) = \sum_j [\phi_j(xyz) \star \delta(x - x_j, y - y_j, z - z_j)], \qquad \ldots\ldots(2.17)$$

where $\phi_j(xyz)$ is the potential distribution which would be obtained if the jth atom were isolated and placed at the origin of coordinates. Taking the Fourier transform of (2.17)

$$E(hkl) = \sum_j g_j(\theta_{hkl}) \exp\{-2\pi i(hx_j + ky_j + lz_j)\}, \qquad \ldots\ldots(2.18)$$

where θ_{hkl} is the Bragg angle for the hkl reflection. Values for the atomic scattering factors g_j, derived by methods which will be discussed below, have been tabulated for use in this connection.

The usual measure of the closeness of agreement between calculated and observed structure amplitudes is the discrepancy factor, or reliability index

$$R = \frac{\Sigma||E_{\text{obs}}|-|E_{\text{calc}}||}{\Sigma|E_{\text{obs}}|}. \qquad \ldots\ldots(2.19)$$

Values of R in the range 0·1–0·2 indicate that the structure deduced is probably correct.

§3. Intensities in Electron Diffraction Patterns

3.1. *Atomic Scattering Factors for Electrons*

For the scattering of electrons through angles in the range normally met with in electron diffraction by crystals, it is probably a sufficiently good approximation to assume that, as in equation (2.2) above, the effect of the potential field of an atom on an incident wave front is given by multiplying by

$$q(xy) = \exp[i\sigma\phi(xy)],$$

where

$$\phi(xy) = \int \phi(xyz)\, dz.$$

For $\sigma\phi(xy) \ll 1$

$$q(xy) \sim 1 + i\sigma\phi(xy)$$

and, by analogy with (2.4), the diffraction pattern is

$$Q(uv) = \delta(uv) + i\sigma g(uv), \qquad \ldots\ldots(3.1)$$

where $g(uv) = g(\theta)$ is the ' kinematic ' atomic scattering factor.

In general, if $\sigma\phi(xy)$ cannot be assumed to be very small

$$Q(uv) = \mathscr{F}\{\exp[i\sigma\phi(xy)]\},$$

where the $\mathscr{F}$ indicates a Fourier transform.

The atomic scattering factor is then complex, being defined by analogy with the kinematic case (3.1) as

$$g'(uv) = \frac{1}{i\sigma}\mathscr{F}\{\exp[i\sigma\phi(xy)] - 1\}$$

$$= g(uv) + \frac{i\sigma}{2}[g(uv) \star g(uv)] - \ldots. \qquad \ldots\ldots(3.2)$$

If as a second approximation only the first two terms of this expression are taken, the atomic scattering factor is seen to be complex with a phase angle which :

(i) increases with atomic number, since the value of $\phi(xy)$ and hence of $g(uv)$ increases with Z ;

(ii) increases as the scattering angle increases, since the convolution $g(uv) \star g(uv)$ will decrease more slowly with angle than $g(uv)$;

(iii) decreases as the accelerating potential of the electrons is increased since σ is proportional to $W^{-1/2}$.

Historically, the atomic scattering factors for electrons have been derived as an extension of the quantum-mechanical treatment of collision processes. The appropriate wave equation has been solved to various degrees of approximation

by various workers. The usual approximation used for electron diffraction work is Born's first approximation, in which all terms representing modification of the potential field by the presence of the electrons are ignored. Then, it is deduced (e.g. Thomson and Cochrane 1939) that the atomic scattering factor is given by

$$g(\theta) = 2\int_0^\infty \phi(r)\left(\frac{\sin sr}{sr}\right) r^2\, dr, \qquad \ldots\ldots(3.3a)$$

where $s = (4\pi \sin\theta)/\lambda$.

Thus the atomic scattering factor is given by the three-dimensional Fourier transform of the radially symmetrical potential distribution.

The potential distribution may be related to the charge distribution by use of Poisson's equation. The charges present in an atom are $+Ze$ in the nucleus and those of the electrons which have a density distribution $\rho(r)$. Consequently the atomic scattering factor may be expressed as

$$g(\theta) = \frac{e}{4\pi}\{Z - f(\theta)\}\Big/\left(\frac{\sin^2\theta}{\lambda^2}\right), \qquad \ldots\ldots(3.3b)$$

where $f(\theta)$ is the x-ray atomic scattering factor, given by Fourier inversion of the electron density distribution in the atom.

Until recently the first Born approximation was considered adequate for all electron diffraction work, and atomic scattering factors for electrons were calculated by the formulae (3.3) or its equivalent using the tabulated values of $f(\theta)$. The first indications that the Born approximation might not be satisfactory came from gas diffraction work on molecules containing both light and heavy atoms. In particular, the array of fluorine atoms in uranium hexafluoride molecules appeared to be rather puzzlingly asymmetric instead of octahedral as might have been expected from other evidence. Schomaker and Glauber (1952) (see also Glauber and Schomaker 1953) showed that the discrepancy could be explained by a failure of the first Born approximation for such systems. Using the second Born approximation, which results in a definition of atomic scattering factor analogous to the first two terms of (3.2), they deduced that the atomic scattering factor for electrons is, in general, complex with a phase factor which varies with atomic number, scattering angle and electron voltage as described above. Good agreement was then obtained between calculated and observed diffraction effects.

The field of applicability of the second Born approximation is rather limited. Hoerni and Ibers (1953) subsequently used the partial-waves scattering theory to obtain atomic scattering factors more generally valid, and have listed magnitudes and phase angles as a function of scattering angle for a selection of elements from $Z = 1$ to 98 for an accelerating voltage of 39·47 kev. They also give transformation formulae for use in calculating the complex atomic scattering factors for other elements and other accelerating voltages.

Insufficient work has been done to allow an evaluation of the importance of the failure of the Born approximation in structure analytical work. The maximum angles of scattering usually employed in crystal diffraction are much less than those used for gas diffraction. The difference in phase angle of the atomic scattering factors for light and heavy atoms are therefore not so great on the

average. Experience with single-crystal patterns of basic lead carbonate (Cowley 1956 a) and with ring patterns of thorium oxide (Hoerni 1957) suggests that the modifications of the intensities may be appreciable. The maximum effect is to be expected when the total scattering from the light atoms is approximately equal in magnitude to the scattering from heavy atoms. Even then the major effect will be confined to weaker spots at relatively high angles of scattering, and in the absence of any heavy atoms in the structure, the real atomic scattering factors of the first Born approximation will continue to be used. These may be

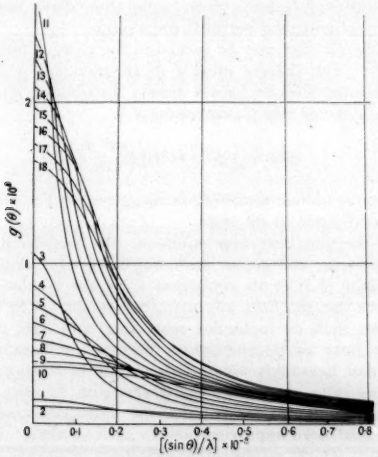


Figure 1. Atom scattering factor curves for electrons for atoms of atomic number 1 to 18. (*After* Weinstein 1953.)

obtained from tabulations given, e.g. by Weinstein (1953), or may be calculated readily from the tabulated x-ray diffraction scattering factors. There are some reservations concerning the use of such atomic scattering factors. Firstly, since the value of the atomic scattering factor for x-rays $f(\theta)$ tends to Z for $\theta \to 0$, the value of $g(0)$ given by equations (3.3) is indeterminate for $\theta = 0$. For θ small, f is very close to Z, and small uncertainties in the $f(\theta)$ values give rise to correspondingly large uncertainties in $g(\theta)$.

Weinstein (1953), using the best available values for $f(\theta)$ at small angles, obtained values of $g(0)$ by extrapolation. His atomic scattering factor curves are shown in figure 1. The behaviour of the curves at small angles is rather surprising. At large angles, for which $f(\theta)$ is small compared with Z, the atomic scattering factor is approximately proportional to Z, as might be expected from (3.3). At small angles and especially for $\theta = 0$, the dependence on Z is reversed within each period of the periodic table of elements. Thus, for $Z = 1, 3, 11, \ldots, g(0)$

is comparatively large, but it decreases with subsequent increase in atomic number within each period. The dependence of $g(0)$ on Z thus parallels the dependence of the atomic radius on Z. More recently Ibers (1957) has shown that $g(0)$ is proportional to the mean square atomic radius. He obtains values for $g(0)$ differing in magnitude from those of Weinstein but having the same sort of dependence on Z.

The value of $g(0)$ corresponds to the integral of $\sigma\phi(xyz)$ over all space for an isolated atom. If a crystal could be considered as made up of an assembly of atoms each unaffected by its neighbours, the sum of the values of $g(0)$ for all atoms in the unit cell divided by the unit cell volume should give the value of the mean potential within the crystal, i.e. the inner potential. Agreement with experiment is good for some non-ionic crystals.

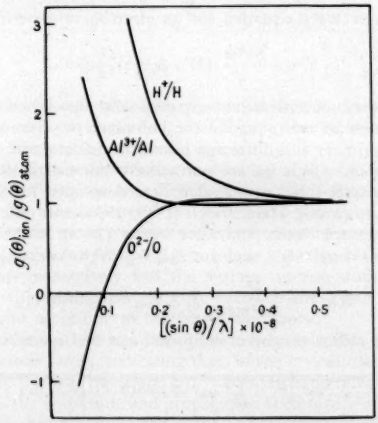


Figure 2. Ratio of atom scattering factors of an ion and its parent neutral atom plotted as a function of $\sin\theta/\lambda$. (*After* Cowley 1953 b.)

The value of $g(0)$ may also be considered as a measure of the ' weight ' of a peak in a potential distribution map obtained by Fourier analysis. As such, it may be possible to obtain $g(0)$ experimentally in the course of structure analysis. For example, Cowley and Ibers (1956) calculated a value for $g(0)$ (actually $f_{el}(0)$ in Å) for chlorine in the course of a study of the structure of a graphite–ferric chloride compound. Their value, $f_{el}(0) \leqslant 6{\cdot}4$ Å, was in agreement with Weinstein's theoretical value of 6·34 Å. (*Note*: Cowley and Ibers overlooked a multiplying constant in quoting Weinstein's value.)

Finally, it must be emphasized that the calculations of atomic scattering factors so far discussed refer only to scattering from isolated, spherically symmetrical, electrically neutral atoms, and must be used with caution, particularly in the

small angle region in dealing with atoms in a crystal lattice which may be ionized or in which the potential distribution may be affected by bonding to neighbouring atoms.

Ionization affects the x-ray atomic scattering factor appreciably in the range of $(\sin\theta)/\lambda$ from 0 to $\sim 0{\cdot}2$. From equations (3.3) it follows that $g(\theta)$ is very much more strongly influenced by ionization in this range and tends to plus or minus infinity as θ tends to zero for positive or negative ions respectively. Plots of $g(\theta)$ for an ion divided by $g(\theta)$ for the corresponding neutral atom, calculated from (3.3), are shown in figure 2. Since the formulae (3.3) cannot be applied to atoms ionized or bonded in a crystal lattice, these results are, at best, only qualitative indications of the changes to be expected on ionization.

3.2. *Diffraction by Perfect Crystals*

The Schrödinger wave equation for an electron in a periodic potential field $\phi(xyz)$ is given by

$$\nabla^2\psi + \frac{8\pi^2 me}{\lambda^2}[W + \phi(xyz)]\psi = 0. \qquad \ldots\ldots(3.4)$$

The dynamic theory of diffraction, as originally developed by Bethe (1928), involves the solution of this equation for boundary conditions corresponding to the behaviour of primary and diffracted beams at incident and exit surfaces of the crystal. The case in which we are particularly interested, that of transmission through a thin parallel-sided crystal slab, was developed in detail by Blackman (1939), MacGillavry (1940), Heidenreich (1949, 1950) and others.

For convenience we denote the wave vector $(2\pi/\lambda)$ of the incident beam in vacuum by $\mathbf{K} = (8\pi^2 emW/h^2)^{1/2}$, and put $E_{hkl} = (Vh^2/8\pi^2 me)v_{hkl}$, so that

$$\phi(xyz) = \frac{h^2}{8\pi^2 me}\sum_h\sum_k\sum_l v_{hkl}\exp\{2\pi i(\mathbf{h}.\mathbf{r})\},$$

where $\mathbf{h}$ and $\mathbf{r}$ are radius vectors in reciprocal and real space respectively. Then (3.4) becomes

$$\nabla^2\psi + \left[\mathbf{K}^2 + \sum_{hkl} v_{hkl}\exp\{2\pi i(\mathbf{h}.\mathbf{r})\}\right]\psi = 0. \qquad \ldots\ldots(3.5)$$

The solution is in the form of a superposition of plane waves

$$\psi = \sum_{hkl}\psi_{hkl}\exp\{i(\mathbf{k}_0 + 2\pi\mathbf{h}.\mathbf{r})\}, \qquad \ldots\ldots(3.6)$$

where $\mathbf{k}_0$ is the wave vector of the incident wave in the crystal.

Putting $\mathbf{k}_h = \mathbf{k}_0 + 2\pi\mathbf{h}$, substituting (3.6) in (3.5) and equating coefficients of exponential terms gives the set of equations

$$\psi_h(K^2 + v_0 - k_h^2) + \sum_g{}' v_g\psi_{h-g} = 0, \qquad \ldots\ldots(3.7)$$

where the triple index hkl has been abbreviated to h. The prime indicates that the zero term v_0 is not included in the summation.

Since, in practice, $v_g \ll K^2$, the amplitude ψ_h of a diffracted beam is appreciable only for small values of $K^2 + v_0 - k_h^2$. This is equivalent to the Bragg condition for a reflection.

If the crystal has a mean potential given by $E_{000} = E_0$, it will have an effective refractive index given by the ratio of the wave vector magnitudes in the crystal (k) and in vacuum (K), i.e.

$$\mu = (K^2 + v_0)^{1/2}/K = k/K = (1 + E_0/W)^{1/2}.$$

The set of equations (3.7) can be solved, except for special cases, only formally. Bethe, Blackman and MacGillavry considered the case in which only two beams, the primary beam ψ_0 and one diffracted beam ψ_h, have appreciable amplitude. The set of equations (3.7) then reduces to

$$\left.\begin{aligned}(k^2 - k_0^2 - U_{00})\psi_0 - U_{0h}\psi_h &= 0\\ -U_{h0}\psi_0 + (k^2 + k_h^2 - U_{hh})\psi_h &= 0\end{aligned}\right\} \quad \ldots\ldots(3.8)$$

where

$$U_{00} = \sum_g \frac{v_g v_{-g}}{k^2 - k_g^2}; \qquad U_{hh} = \sum_g \frac{v_{h-g} v_{g-h}}{k^2 - k_g^2};$$

$$U_{h0} = -v_h + \sum_g \frac{v_{h-g} v_g}{k^2 - k_g^2}; \quad U_{0h} = -v_{-h} + \sum_g \frac{v_{-g} v_{h-g}}{k^2 - k_g^2}.$$

Here U_{00} and U_{hh} may be interpreted as apparent inner potentials for the primary and diffracted beams respectively, and U_{0h} and U_{h0} may be called dynamic structure amplitudes corresponding to the kinematic v_h and v_{-h}.

The boundary conditions which must be satisfied at the crystal–vacuum interface are that the wave functions and their first derivatives should be continuous, which implies that the components of the wave vectors tangential to the surface should be equal. Hence there are relations which must be satisfied between the magnitudes of the wave vectors $\mathbf{k}$ and $\mathbf{K}$, the angles χ and χ_h between the primary and diffracted beams respectively and the surface normal, and ζ_h the so-called 'excitation error' representing the distance between the Ewald sphere and the reciprocal lattice point according to the kinematic theory.

Applying these relations, the equations (3.8) may be solved to give two solutions for ψ_0 and two solutions for ψ_h, indicating that, in the crystal, there are two waves corresponding to the primary beam and two waves corresponding to the diffracted beam. In each case the amplitude and propagation vectors of the two waves differ.

For transmission through a parallel-sided slab of thickness H, the boundary conditions on the exit face are identical with those of the entrance face except for the reversal of sign of the surface normal vector. The two crystal waves corresponding to the diffracted beam will have, on entering the vacuum, the same propagation vector, and their amplitudes will therefore be added together to give the diffracted wave observed. Because there is a difference in propagation vector in the crystal, the two waves emerge with a phase difference proportional to H and the magnitude of the resultant vacuum wave Ψ_h therefore varies sinusoidally with H with a periodicity dependent on the quantities ζ_h and v_h, which both influence the propagation vectors within the crystal.

If we made the approximations that the summations over g in the definitions of U_{00}, U_{hh}, U_{0h} and U_{h0} in (3.8) may be neglected and that the angles of the beams with the face normals are very nearly right angles, and putting

$$\zeta_h = k(\theta - \theta_0)\sin 2\theta_0$$

where θ is the angle between the primary beam and the diffracting plane and θ_0 is the Bragg angle for the exact fulfilment of the kinematic conditions for reflection, the amplitude of the diffracted wave is given by

$$\left|\frac{\Psi_h}{\Psi}\right| = \frac{v_h \sin\{\frac{1}{2}kH[(\theta-\theta_0)^2\sin^2 2\theta_0 + v_h^2/k^4]^{1/2}\}}{k^2[(\theta-\theta_0)^2\sin^2 2\theta_0 + v_h^2/k^4]^{1/2}}, \qquad \ldots\ldots(3.9)$$

where Ψ is the incident beam amplitude.

The kinematic theory expression may be derived from (3.9) by letting v_h tend to zero, i.e. by assuming that the propagation of the waves in the crystal is not influenced by the crystal structure. Then

$$\left|\frac{\Psi_h}{\Psi}\right| = \frac{1}{2}\frac{v_h}{k}H\frac{\sin\{\frac{1}{2}kH(\theta-\theta_0)\sin 2\theta_0\}}{\{\frac{1}{2}kH(\theta-\theta_0)\sin 2\theta_0\}}. \qquad \ldots\ldots(3.10)$$

This may be interpreted as the amplitude given when the Ewald sphere cuts the shape transform of the parallel-sided crystal at a distance ζ_h from the reciprocal lattice point in a direction parallel to the incident beam, since the shape transform for such a crystal will be of the form

$$\{\sin(\tfrac{1}{2}H\zeta_h)\}/(\tfrac{1}{2}H\zeta_h).$$

By analogy equation (3.9) may be interpreted in terms of the intersection of the Ewald sphere with a ' dynamic shape transform ' differing from the kinematic shape transform in that it is a function of v_h and so is not the same for all reciprocal lattice points and also varies with the electron wavelength.

All the attempts to confirm experimentally the predictions of the dynamic theory regarding the intensities of reflections have been concerned with patterns from polycrystalline material in which there is a considerable range of crystal thicknesses involved and all possible orientations of the crystal contribute equally to the intensity. The effect of thickness variation is given by averaging the square of (3.9) over all values of H. Since the average value of $\sin^2 x$ is $\frac{1}{2}$, we have

$$\overline{I_h/I_0} = v_h^2/2(k^2\zeta_h^2 + v_h^2).$$

If we now sum the intensity for all values of ζ_h, and so for all crystal orientations, we obtain the integrated intensity for the reflection

$$I_h \text{ (integrated)} = I_0\int_{-\infty}^{+\infty}\frac{v_h^2}{2(k^2\zeta_h^2 + v_h^2)}d\zeta_h \propto \frac{I_0\pi v_h}{2k^2} \propto E_{hkl}. \qquad \ldots\ldots(3.11)$$

If we carry out the same averaging process for the kinematic theory result, (3.10), the integral of (3.11) does not contain v_h in the denominator. Hence we obtain

$$\overline{I_h} \text{ (integrated)} \propto v_h^2 \propto E_{hkl}^2. \qquad \ldots\ldots(3.12)$$

Intensities of diffraction rings therefore depend on the first power of the structure factor according to the dynamic theory and on the second power according to the kinematic theory. However, if the range of thicknesses H involved is small this averaging process is not valid. If, for example, H is so small that the argument of the sine in (3.9) is much less than π, the expression [sin (arg)]/[arg] is close to unity, and the integrated intensity will consequently be proportional to E_{hkl}^2 as in the kinematic theory. There is thus a range for the average crystal

thickness below some critical thickness, say H_c, in which the dynamic and kinematic theories give the same dependence of the intensity on the structure factor. Blackman's (1939) estimates of this critical thickness for the strongest reflections of metals of simple structure, and Weinstein's (1955 a) estimates for average reflections from more complicated structure have been discussed in § 1 above.

The extension of the dynamic theory to cases in which more than two beams have appreciable amplitude in the crystal has been considered by various authors; notably by Heidenreich (1950) in his investigation of the origin of the ' forbidden ' 222 reflections in the diamond structure. If the strong beams are taken as the primary beam ψ_0 and two diffracted beams ψ_h and ψ_s, then the set of equations (3.7) is reduced to four equations, namely, the equations for ψ_0, ψ_h and ψ_s and the equation for ψ_λ, where $\lambda = h - s$, which must be included since the summation in the second part of the equation contains terms in ψ_s and ψ_h. When the boundary conditions are applied, these four equations lead to four solutions for the primary wave and four for each diffracted wave. There is thus, in effect, quadruple refraction. Heidenreich showed that, under these conditions, the intensity of, for example, the s diffracted beam may be as great as or greater than that of the other diffracted beams even if ψ_s corresponds to a reflection ' forbidden ' under kinematic diffraction conditions, such as the 222 reflection from a diamond-type lattice. The modification of ψ_s by the presence of other diffracted beams (h and λ not forbidden, such that $s = h - \lambda$) may be attributed to ' dynamic interaction ', or the coherent addition of the singly and multiply reflected waves within the crystal.

An extrapolation from the three-beam case to the n-beam case would suggest that the effect of dynamic interaction on an extensive spot pattern given by a single crystal which is not too thick would be to smooth out the variations in intensity between diffraction spots. In particular, spots forbidden by space-group considerations for the kinematic theory will appear with appreciable intensity if the indices of the forbidden spots can be obtained by addition or subtraction of the indices of ' allowed ' spots. Such an extrapolation process can, however, be misleading. The treatment of dynamic scattering in the single-crystal spot patterns containing a multitude of spots is best approached from an entirely different viewpoint (Cowley and Moodie 1957 d, 1958). Here the problem is treated by first considering the multi-spot pattern given by a very thin crystal and then determining through a series of converging approximations, the modifications of the pattern consequent on an increase in crystal thickness.

For extremely thin crystals the effect on the electron beam may be represented by the kinematic approximation for very small wavelength, given by equation (2.3). As the next higher approximation, valid for thicker crystals, the phase-grating nature of the crystal lattice for an electron beam must be taken into account by considering the effect on the electron beam to be given by (2.2). This introduces modification of the intensities due to a phase change in the atomic scattering factor for each atom proportional to the thickness of the crystal. Since, however, the symmetry of $\exp\{i\sigma\phi(xy)\}$ is the same as that of $\phi(xy)$, there are forbidden reflections as for the kinematic theory. The next higher approximation, consisting of dividing the crystal into two phase gratings, each of half the total thickness, introduces the possibility of double scattering of diffracted beams but no variation

of the kinematic specification of forbidden reflections and similarly for higher approximations involving multiple scattering processes. It follows that, if the electron beam is parallel to a principal axis of the crystal, the diffraction pattern retains, for all crystal thicknesses, the symmetry of the reciprocal lattice perpendicular to the beam and the relevant space-group absences. In the single-crystal patterns of greatest interest, therefore, dynamic scattering cannot give rise to forbidden reflections.

Another conclusion which may be drawn is that, if the effective phase change of the atomic scattering factor as a function of thickness is appreciated, the presence of some ' dynamic ' effects, or a failure of the kinematic approximation, may not prevent the use of the spot intensities for purposes of structure analysis.

3.3. *Diffraction by Imperfect Crystals*

The results obtained for the diffraction of electrons from perfect crystals must be applied with considerable care to the consideration of diffraction from real crystals, since only in exceptional cases can the definition of a perfect crystal be taken as a reasonable approximation to a diffracting crystal in the size range dealt with in electron diffraction transmission experiments. Substances crystallize in this range of sizes, in general, either because the crystals are grown quickly under conditions of excess energy, or because, as in the case of some clay minerals, there is a limitation to the size to which crystals may grow as a result, presumably, of an accumulation of defects or strains in the lattice. In either case, it is to be expected that the resultant crystals will have a higher density of lattice defects than is usual in the larger crystals, such as those employed for x-ray diffraction work.

Apart from defects on an atomic scale, such as vacant lattice sites or interstitial atoms, which will not normally have an appreciable effect on the diffraction pattern unless in very high concentrations, the most common defects will be defects in the stacking order of lattice planes and the lattice distortions surrounding dislocation lines. ' Stacking ' defects are particularly common in the layer–lattice structures, which have constituted a high proportion of the structures investigated by electron diffraction methods and will probably continue to do so because of the relative ease of preparing suitable specimens of these substances.

The diffraction of x-rays by crystals containing stacking faults has been studied by a number of authors. A general survey of the subject is given in the book by James (1948). More detailed treatments have been given in papers by Hendricks and Teller (1942), Jagodzinski (1949 a, b), Kakinoki and Komura (1952) and Gevers (1954) among others.

One of the types of stacking fault most commonly treated is that which occurs in those structures which can be considered as built up by the closest packing of spherical atoms, or groups of atoms. The individual layers in this case are hexagonal nets of atoms, with a unit cell as shown in figure 3, with atoms in the positions A. The layer above that shown may have its one atom per unit cell directly above either the point B or the point C. The inter-layer energy is the same in the two cases. The layer beyond a B layer can be either a C or an A layer, and so on. The actual stacking sequence is determined by forces between next-nearest layers. The most common sequences are of the ABABAB . . . type, corresponding to hexagonal close-packing, or of the ABCABCABC . . .

type, with a face-centred cubic unit cell. Under some circumstances the energy difference between these two stacking sequences is very small and mistakes or changes from one sequence to the other are common. An extreme case, for which it is relatively easy to calculate the diffraction effects, is that for which the stacking sequence is random among the positions A, B, C, with the exception that an A layer cannot follow an A layer, and similarly for B and C layers. The diffraction effects may be described relative to a unit cell with the hexagonal cell of figure 3, defined by the a and b axes, and the c axis perpendicular to this layer plane and of length equal to the inter-layer separation. In reciprocal lattice space then, there are no longer sharp maxima of scattering power at all points hkl,

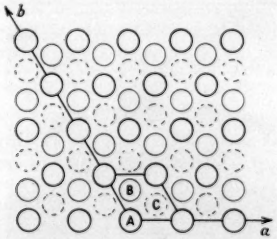


Figure 3. Possible stacking arrangements of planes of close-packed atoms.

but only for those points for which $h+k=3n$. For $h+k\neq 3n$ the rows of spots parallel to the c^* axis having the same h and k values are replaced by continuous lines of scattering power. In close-packed systems having a smaller degree of randomness, these continuous lines also occur, but with a correspondingly different distribution of scattering power along their lengths.

In electron diffraction experiments, the intersection of the Ewald sphere with such continuous lines can give rise to the appearance of anomalous ' one-third order ' spots, or, if the Ewald sphere very nearly contains parts of the continuous lines, to sharp lines or streaks in the diffraction pattern. Examples which may be mentioned include the streaks and spots observed in patterns given by electrolytically deposited silver films, and other metal films (see, for example, Gottsche 1956), and the somewhat similar effects found by Cowley (1953 a) in patterns given by small single crystals of γ–Al_2O_3 (figure 4).

When the individual single crystals in a polycrystalline specimen contain such closest-packing stacking faults the hkl reflections give sharp rings for $h+k=3n$. For $h+k\neq 3n$ the set of rings corresponding to all the hkl reflections for a given h and k is replaced by a broad band with a sharply defined inner edge and an intensity which decreases, possibly through a series of maxima and minima, as the scattering angle increases.

For the close-packed structures, the inter-layer displacements are limited to those given by the vectors AB and AC in figure 3. In other systems the possible inter-layer displacements may not be simply related to the unit cell dimensions or may have a random distribution. Then all rows of spots parallel to the c^* axis

except that through the origin will be replaced by continuous lines of scattering power. The calculations of the x-ray diffraction effects from crystals with disordered layer stacking make the assumption that the number of stacking faults in a crystal is very large. Under these conditions it is justifiable to assume that, if faults with random inter-layer displacements occur, the regions of the crystal between such faults can be considered as diffracting independently, and the diffraction intensities may be assumed equal to the intensities given by an average region.

Figure 4. Continuous lines, arising from stacking faults, passing through a 004 spot in a γ-alumina pattern. Beam parallel to [110]. (*After* Cowley 1953 a.)

Only in extreme cases can such assumptions be made in single-crystal electron diffraction. A crystal to which the kinematic theory of diffraction may be applied will usually contain no more than a few hundred layers in the direction of the incident beam. Unless stacking faults occur with average separations equivalent to a very few layer spacings, the number of faults occurring will not be great enough to justify the use of statistical methods. On the other hand, in patterns given by a large number of crystals each with the same type of stacking defects, such as patterns from mosaic single crystals, and arc or ring patterns from a very large number of crystals, the number of unrelated crystal regions contributing to any reflection is usually great enough, and the statistical methods may be used.

In either case, the diffraction intensities may be calculated or observed intensities may be interpreted by employing the concept of a distribution function (Cowley 1957). The potential distribution in a crystal may be expressed as

$$\phi(xyz) = \phi^0(xyz) \star D(xyz), \qquad \ldots\ldots(3.13)$$

where $\phi^0(xyz)$ is the potential distribution of a single layer (or of any other suitable sub-division of the crystal) and the distribution function, $D(xyz)$, is a set of δ-functions in the positions to which a given origin point in $\phi^0(xyz)$ must be translated in order to build up the total potential distribution $\phi(xyz)$.

Taking the Fourier transform of both sides of (3.13) gives the reciprocal space relationship

$$E_{hkl} = C_{hkl} E_{hkl}^0, \qquad \ldots\ldots(3.14)$$

where E_{hkl} and E_{hkl}^0 are the structure factors for the whole crystal and the individual layers respectively and C_{hkl} is the value of the Fourier transform of $D(xyz)$ at the point hkl. Since $D(xyz)$ is not, in general, centro-symmetric, C_{hkl} is generally complex and it is not usually practicable to use the equations (3.13) and (3.14) for purposes of calculation or interpretation. Instead we consider the corresponding equations for quantities derivable from intensity data, namely,

$$E_{hkl} E_{hkl}^* = (C_{hkl} C_{hkl}^*)(E_{hkl}^0 E_{hkl}^{0*}) \qquad \ldots\ldots(3.15)$$

and

$$P(XYZ) = Q(XYZ) \star P^0(XYZ), \qquad \ldots\ldots(3.16)$$

where $P(XYZ)$ and $P^0(XYZ)$ are the Patterson functions for the whole crystal and the individual layer, and $Q(XYZ)$ is the Patterson function corresponding to the distribution function. $Q(XYZ)$ will give the magnitude and direction of displacement vectors between lattice origins and the relative frequencies with which they occur. This is all the information which is normally required concerning the stacking of the layers in the crystal. Alternatively, $Q(XYZ)$ can be considered as an average distribution function, since it can be made up by taking all the distribution functions referred to origins in individual layers and superimposing them with a common origin.

If the structure of the individual layers is known, equations (3.15) and (3.16) may be used to determine the stacking arrangement of the layers within the crystal from the observed intensities. If the layer structure is not known, but the distribution function is partially known, these equations may be used as a basis for a refinement process leading to a determination of both the layer structure and the distribution function. This was done in the course of the structure analysis of basic lead carbonate from single-crystal electron diffraction data, when the two-dimensional analogues of equations (3.13) to (3.16) were used (Cowley 1956 a).

Of the many other types of structural imperfection which may occur in crystals giving electron diffraction patterns, the x-ray diffraction effects of most have been calculated, and it appears likely that the electron diffraction effects can be assumed to be given adequately by the same theoretical treatment at least to the degree of approximation involved. Some mention has already been made of the intensity anomalies occurring in patterns from thin spine-like crystals, such as those of ZnO and CuO′, which contain screw dislocations parallel to the spine axis. The predictions of Wilson (1952) concerning the diffraction effects to be expected from crystals containing screw dislocations were shown to be at least qualitatively correct in these studies.

A phenomenon which is of much greater importance in electron diffraction than in x-ray diffraction work is ' secondary elastic scattering '. In our discussion of diffraction by crystals with disordered layer stacking we treated diffraction by

successive regions, even if separated by arbitrary displacements, as coherent; that is, the amplitude of the diffracted beam is given by adding the amplitudes from all crystal regions. This is certainly a correct procedure under all experimental conditions which we will be considering. If, however, the individual crystal regions do not have the same orientation, one being rotated or tilted with respect to another, or if they are different in structure, having differing unit cell dimensions, the diffraction spots from the various regions will not coincide and, if overlapping of the individual spots does not occur, the spot intensities obtained by summing the amplitudes from all regions will be the same as the intensities from each region individually. One may then consider the crystal regions as diffracting independently. If electrons are diffracted successively by two crystals, all such double-diffraction processes can be considered separately. The diffraction pattern will then contain, apart from the primary spots due to the two crystals taken separately, secondary spots in positions given by making each spot from the first crystal in turn the origin of a spot pattern of the second crystal. The intensity of a secondary spot is proportional to the product of the intensities of the contributing primary spots.

The most comprehensive account of secondary elastic scattering effects is probably that given by Cowley, Rees and Spink (1951 a), who observed a great variety of such effects from multiple crystals of the long chain n-paraffin dicetyl $C_{32}H_{66}$. The experimental observations explained in this way included large numbers of spots with irrational indices, rings of spots surrounding strong single-crystal spots, well-defined envelopes of rings of secondary origin which may be mistaken for extra primary rings, and small regular arrays of spots surrounding primary single-crystal spots. In some cases, apparently, the crystal regions concerned were so nearly parallel that they appeared to give a single spot pattern, the spots from the separate regions being unresolved because of the finite source size. Under these conditions the secondary spots would also fall very close to the primary spot positions and would not be resolved from them. The appearance was thus that of a single-crystal pattern with spots only a little larger than the primary beam spot size, but with spot intensities modified by the addition of the secondary scattering intensities. In particular, the $h00$ and $0k0$ spots with h or k odd, which have zero intensity in patterns from thin perfect crystals, being forbidden by the symmetry of the structure, appeared with appreciable intensity as a result of secondary diffraction. Such forbidden reflections can be given by successive reflection from two planes whose indices add up to those of the forbidden plane. Thus the 010 reflection of dicetyl may be given by reflection from (210) and ($\bar{2}$00) planes and similar pairs. Methods were evolved whereby the intensities of apparent single-crystal patterns could be corrected for the presence of small amounts of secondary scattering, provided that the presence of reflections which should be forbidden is recognized.

So far we have considered only crystal imperfections for which the description in terms of a unit cell can be maintained, since the operation which makes one set of atoms coincide with another set is one of translation only. This is not the case if the crystal as a whole, or small regions of it, are bent or otherwise distorted. Blackman (1951 a, b) has made a study of the diffraction effects for the simplest system of this sort; a thin crystal plate bent into portion of a circular cylinder

with an axis perpendicular, or nearly so, to the electron beam. This case is of particular importance because of the observation that thin crystal laminae viewed in electron microscopes frequently appear to be bent. Blackman's conclusion was that under normal conditions of diffraction it would not be easy to distinguish the pattern from such a bent crystal from that of a mosaic of small unbent crystalline blocks with a distribution in orientation approaching that of the portions of the bent crystal. More detailed study of the effects of distortion on diffraction patterns is a very difficult matter, both theoretically and experimentally, and very little has been attempted along these lines.

§4. Structure Analysis from Intensities

4.1. *Collection of Intensity Data*

From our discussion of the diffraction of electrons by perfect and imperfect crystals there has emerged a fairly clear picture of the limitations in the dimensions of crystals from which useful intensity data can be obtained. The perfect crystal regions present must not be more than a few hundred ångström units in thickness in the beam direction if the kinematic theory of diffraction, or a pseudo-kinematic theory which takes account of unequal phase changes in diffraction, is to be applied. For imperfect crystals the allowable thicknesses are greater, but not by a very large factor if secondary scattering effects are to be avoided.

The preparation of suitable specimens is governed by imperfectly understood factors which influence the size and habit of the crystals formed under any particular conditions. No general prescriptions for specimen preparation can be given. One of the most productive methods, however, is that of evaporating a small quantity of a dilute solution of the substance under controlled conditions on a suitable base, such as a carrier film of collodion, carbon or any other of the substances used for specimen supports in electron microscopy. Alternatively, a drop or thin film of the solution lying across small holes in a fine metal grid or a thin foil may be evaporated. Substances which volatilize without decomposing may be evaporated in vacuum on to a suitable base. Other substances may best be prepared by chemical reaction carried out *in situ* on a supporting film or grid.

For ring patterns from polycrystalline material, the principal requirements for purposes of intensity measurement are that there should be no preferred orientation of the crystallites and no pronounced correlation of the crystal morphology and lattice structures. Preferred orientation of the crystallites has the effect of producing variations in the intensity along the circumference of the diffraction rings except when one crystal axis is preferentially oriented in the direction of the incident beam. Confusion due to this exception is usually avoided by recording several patterns from specimens tilted through different angles. Any pronounced morphological feature of the crystal which is likely to modify intensities by anisotropic absorption or dynamic effects can normally be detected by electron microscopy or by a study of high-resolution diffraction patterns.

The integrated intensity of a small portion of a powder pattern ring of length Δ is given by Weinstein (1956 b) as

$$I_{hkl}' = I_0 \lambda^2 \left| \frac{E_{hkl}}{V} \right|^2 \frac{U d_{hkl}^2 p \Delta}{4\pi L \lambda}, \qquad \ldots\ldots(4.1)$$

where I_0 is the incident beam intensity, U the volume of the specimen, L the distance from the specimen to the plane of observation and p is the multiplicity factor, which is a measure of the number of planes having the same d_{hkl} value contributing to the ring intensity. From measured intensities on an arbitrary scale relative values for the structure amplitudes may thus be found from the relation

$$|E_{hkl}| \propto (I_{hkl}/d_{hkl}{}^2 p)^{1/2}. \qquad \ldots\ldots(4.2)$$

The use of ring pattern intensities for purposes of structure analysis is severely limited by the superposition or overlapping of rings. The relative contributions of reflections from planes which have the same d_{hkl} spacings, but which are not crystallographically equivalent, cannot be determined separately. The requirement that all the crystal dimensions should preferably be very small implies that the rings will be broadened to the point at which there will be considerable overlapping of rings for all but the simplest of unit cells.

The situation in these respects is considerably improved, although still not without difficulties, if arc patterns from polycrystalline samples with a high degree of preferred orientation are employed. The types of preferred orientation most frequently encountered are those in which one particular lattice plane tends to be parallel to a fixed plane, such as the plane of a supporting film, or a particular lattice axis tends to be oriented perpendicular to a given plane or parallel to a given direction.

Suitable specimens are commonly prepared by taking advantage of special features of the crystal morphology or method of growth. For example, if the crystals have the form of thin flakes or rods they will frequently lie on a neutral substrate with their larger dimensions parallel to its plane.

For good arc patterns with reliable intensity distributions, the crystals should vary in orientation of the preferentially oriented axis or plane by no more than a few degrees on the average, but there should be complete randomness in orientation otherwise, i.e. in rotation about the preferred axes or parallel to the preferred plane. We may consider, for example, an orthorhombic lattice with the c axis of all crystals very nearly parallel to a given direction, and the incident electron beam making an angle χ with this direction. The diffraction pattern will then consist of a number of lines of arcs, known as layer lines, spaced at intervals of $c^*/\sin\chi$ in the direction of tilt of the incident beam. Each layer line contains arcs corresponding to reflections with a given l index; the zero-layer line which passes through the central spot containing $hk0$ reflections, those on either side containing $hk1$ and $hk\bar{1}$ reflections, and so on. An example of a layer-line pattern is given in figure 5.

If the arcs are short, it is convenient to compare the integrated intensities of the arcs, calculated by Weinstein (1956 b) to be given by

$$I_{hkl} = I_0 \lambda^2 \left|\frac{E_{hkl}}{V}\right|^2 \frac{UL\lambda p}{2\pi R' \cos\chi}, \qquad \ldots\ldots(4.3)$$

where R' is the distance of the arc from the centre of the layer line. Structure amplitudes are obtained from the relation

$$|E_{hkl}|^2 \propto (I_{hkl} R'/p)^{1/2}. \qquad \ldots\ldots(4.4)$$

For relatively long arcs the easiest intensity measurement to make is that of the peak intensity at the centre of the arc given by

$$I_{hkl}' = I_0 \lambda^2 \left| \frac{E_{hkl}}{V} \right|^2 \frac{U d_{hkl} d_{hk0}}{\alpha 2\pi L \lambda}, \qquad \ldots\ldots(4.5)$$

where α is a measure of the spread of orientations about the preferred direction.

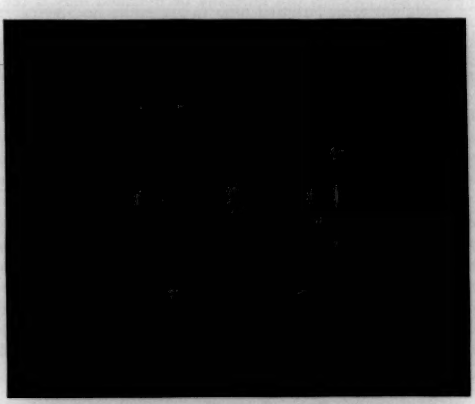


Figure 5. Layer-line pattern of arcs from an inclined and oriented polycrystalline specimen of $MnCl_2.2H_2O$. (*After* Weinstein 1952.)

Not all *hkl* reflections can appear in an arc pattern. Reflections corresponding to reciprocal lattice points in a cone of half-angle $(90-\chi)$ about the c^* axis will be excluded. In our example then 001 reflections will always be excluded unless $\chi = 90°$, which is possible, for example, if the crystals are in the form of thin needles with a preferred orientation of the needle axis in direction at right angles to the beam, provided that it is a justifiable assumption to take the Ewald sphere as planar. Arc patterns are limited also since for some space groups arcs corresponding to reflections from non-equivalent planes will sometimes be superimposed.

If polycrystalline specimens have two degrees of preferred orientation, so that two crystal axes have preferred direction, the patterns are similar to those of mosaic single crystals with a variation in orientation of the mosaic blocks, i.e. regular spot patterns with each spot replaced by a short arc. Specimens giving such patterns may be produced under favourable circumstances as, for example, when the crystals are deposited or are grown in contact with a substrate which is a single crystal.

The intensity of the arcs in such patterns has been shown by Weinstein (1956 a) to be given by

$$I_{hkl} = I_0 \lambda^2 \left| \frac{E_{hkl}}{V} \right|^2 \frac{U d_{hkl}}{\alpha}, \qquad \ldots\ldots(4.6)$$

where α is a measure of the spread in crystal orientations. The structure amplitude is then obtained from the relation

$$|E_{hkl}| \propto (I_{hkl}/d_{hkl})^{1/2}. \qquad \ldots\ldots(4.7)$$

The $1/d_{hkl}$ in this equation may be regarded as a Lorenz factor for electrons for mosaic single crystals.

True single-crystal patterns can normally be obtained only if the area of the specimen examined is reduced to a very small value, since it is only rarely that a crystal of only a few hundred ångström units thickness will extend for more than a micron or two laterally without being bent or having a mosaic sub-structure.

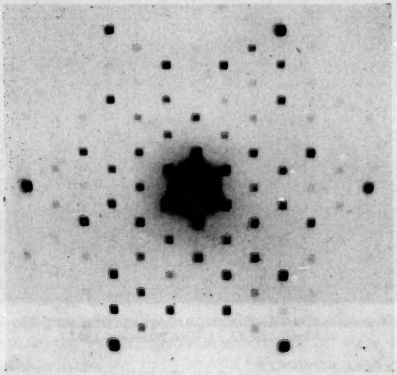


Figure 6. Pattern from disordered single crystal of boric acid demonstrating the technique of spreading each spot over small rectangular area by use of electrostatic deflection. (*After* Cowley 1953 c.)

Very small electron beam diameters may be achieved by the use of precision electron-optical components. The beam may be focused on the specimen provided that the divergence is so small that the resolution of the diffraction pattern is not seriously affected. Single-crystal specimens are most readily prepared when the crystals have naturally a flaky habit and grow readily in the form of thin plates. Hence most single-crystal studies to date have been concerned with layer-lattice structures. A reduction of the specimen area examined should facilitate extension of the technique to other types of crystals.

Apart from specimen preparation, the chief experimental difficulty in the use of single-crystal patterns is that of specimen manipulation. For stationary crystal patterns it is necessary to control the tilt of a crystal with an accuracy of a fraction of a degree in order to get symmetrical intensity distributions in patterns of spots corresponding to principal axis projections. Lateral movements of the specimen of the order of microns must also be possible. Electron microscope viewing of the specimen may be of great assistance in the selection of suitable crystals and their manipulation.

The usual method of recording intensities is photographic. If, as is usual, the range of intensities to be measured is greater than the linear range of the photographic emulsion, a series of exposures is made with exposure times increasing in a geometric progression. Measurement of relative intensities is then done by means of visual inter-comparisons with a calibration given by the multiple exposures or by the use of a microphotometer. It is, however, difficult to use a microphotometer effectively on either arc or spot patterns since the shape and inclination of arcs varies along each layer line, and single-crystal spots are usually small and often of uneven intensity distribution. The difficulty can be overcome for single-crystal spot patterns by using electrostatic deflector plates to spread each diffraction spot over a small rectangular portion of the plate. If the rectangles are of more than twice the dimensions of the diffraction spot, the intensities at the centre of the rectangles will be proportional to the integrated spot intensity. A diffraction pattern with the spots spread in this way is shown in figure 6.

Some experiments have been made on the use of other than photographic methods for the recording of intensities. Lennander (1954) used a thin-windowed Geiger counter to measure the intensity distribution along the radius of a ring pattern. Takagi and Suzuki (1955) made use of the bombardment conductivity induced by electrons in cadmium sulphide to make similar measurements. In both of these cases, although the method appeared to give quite accurate results, the procedure was tedious and the time taken to measure even the intensity distribution along a radial line was so great that general application of the method seems impracticable until the techniques are developed a great deal further.

4.2. *Interpretation of Electron Diffraction Fourier Maps*

A successful structure analysis using Fourier series techniques using the intensity data obtained from electron diffraction patterns gives the distribution of potential in the crystal lattice or, more usually, one or more two-dimensional projections of the potential distribution. A potential distribution map, while resembling the corresponding electron density map obtained by x-ray diffraction methods, differs from it in several important respects. The positions of the maxima will generally coincide for atoms of moderate atomic weight, but for ions of the lighter atoms, and especially hydrogen, placed in a non-symmetrical field, the position of maximum electron density may not coincide exactly with the position of maximum potential, since both the electrons and the positively charged nuclei contribute to the potential distribution. Potential maps may not appear so satisfying to an eye accustomed to judge electron density maps since the atomic peaks may not be so sharply defined, and the background between peaks is usually not so flat, having more definite minima in spaces between atoms.

A number of techniques are available whereby electron density maps may be placed on an absolute basis. Commonly the criterion is used that the ratio of calculated to observed intensities, plotted as a function of scattering angle, should tend to unity as the scattering angle tends to zero. Because of the uncertainty of the atomic scattering factors for electrons, especially for small scattering angles, this method is of little use in electron diffraction work, and it is better simply to scale the observed intensities to give best agreement with the calculated intensities for the medium scattering angle range. In particular cases it may be possible to

obtain absolute values for one or more structure amplitudes from the fine-structure measurements to be described in § 5. These could then be used to put all other structure amplitudes on an absolute basis.

In the discussion of the value of $g(\theta)$ as given by the first Born approximation to the atomic scattering factor, it was mentioned that $g(0)$ gave a measure of the integrated potential due to an isolated atom. It is not to be expected, however, that the height of peaks in potential maps would depend very much on $g(0)$. The peak height will depend more on the values of $g(\theta)$ at relatively high scattering angles since it is intensities in this region which determine the sharpness of the peak. Since $g(\theta)$ in this region increases monotonically with the atomic number Z, the peak height may be expected to do likewise.

The state of ionization of an atom may have some influence on the relative peak heights, since for unit cells of a moderate size there are usually several reflections within the region of scattering angles for which ionization has an appreciable effect. These subjects are discussed in greater detail in the book by Weinstein (1956 c).

4.3. *Structure Analysis from Polycrystalline Material*

The first application of Fourier series techniques to the determination of unknown crystal structures from electron diffraction data was the study of the structures of $BaCl_2$ and $BaCl_2.H_2O$ made by Weinstein and Pinsker (1949). Similar studies were subsequently made of the other hydrated compounds $CoCl_2.2H_2O$ (Weinstein 1949), $BaBr_2.H_2O$ (Weinstein and Pinsker 1950) and $MnCl_2.2H_2O$ (Weinstein 1952). The intensities were measured from patterns of arcs on layer lines due to polycrystalline material with a high degree of preferred orientation about an axis making a considerable angle to the electron beam. Such a pattern, given by $MnCl_2.2H_2O$ is shown in figure 5. Patterns from mosaic single crystals were also observed.

Patterson projections were successfully interpreted and two-dimensional projections of the potential distributions were calculated. Figure 7 shows the projection along the c axis of the structure of $BaCl_2.H_2O$.

The Russian workers have since developed their techniques of both observation and interpretation of the intensity data from arc patterns with an occasional subsidiary use of single-crystal patterns from mosaic crystals or of ring patterns from samples with no preferred orientations. They have performed successful structure analyses of a wide range of compounds, including studies of metal complex formation in $KPtCl_2NH_3$ and Cs_2CoCl_4, various carbides and nitrides of iron, selenides and tellurides of bismuth and several organic compounds (Weinstein 1956 c).

The structure analysis of paraffin and of diketopiperazine are of particular interest in that they demonstrate the power of the electron diffraction method in the detection and location of hydrogen atoms and ions in organic structures. In each case the positions of the heavier atoms were known from x-ray diffraction work, but the electron density maps obtained from the x-ray data had not given any more than very vague indications of the presence of hydrogen atoms.

Layer-line patterns of arcs from oriented polycrystalline specimens were employed in the study of paraffin wax (Weinstein and Pinsker 1950) in which

44 *hk*0 reflections and 47 *hkl* intensities were measured by means of the multiple exposure technique. From the *hk*0 intensities a two-dimensional projection on the plane *ab* was calculated, in which well resolved small peaks indicated the hydrogen positions. To obtain a more accurate picture of the CH_2 group, the *hkl* reflections were incorporated in a three-dimensional Fourier series summation to give the two-dimensional section of the lattice in the central plane of the

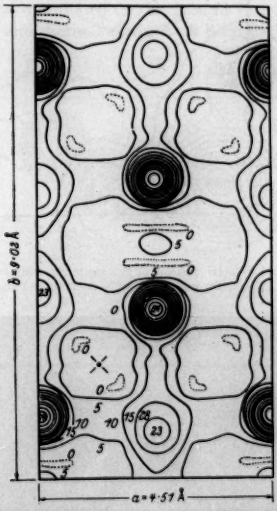


Figure 7. Projection of potential distribution of $BaCl_2.H_2O$ along the *c* axis. (*After* Weinstein and Pinsker 1949.)

group, shown in figure 8. From this it is concluded that the C–H distance is 1·17 Å and the H–C–H angle is 105°. The C–H bond length found is appreciably greater than the 1·07 to 1·09 Å expected from spectroscopic data. The relative peak heights for carbon and hydrogen were found to be in the ratio 3·5 to 1, which is considerably different from the ratio of atomic numbers, 6 to 1. The difference is presumably due to the increase in the atomic scattering factor of hydrogen due to a partial ionization.

The study of diketopiperazine (Weinstein 1955 b) is probably the most thorough and complete analysis of a structure yet attempted by electron diffraction methods. From arc patterns such as figure 9, the intensities of 280 reflections were measured with an estimated accuracy of 30%, i.e. an accuracy of 15% in structure amplitudes. In 25 cases, mostly for faint reflections, two arcs overlapped and could not be separated. The intensities for these cases were divided in the ratio of the calculated intensities. The 0*k*0 reflections could not be observed since the

b axis was the axis of preferred orientation, and estimated intensities were used. Projections on the *ac* and *ab* faces were calculated. The hydrogen atoms gave rise only to bumps on the C and N peaks in these projections. In order to locate

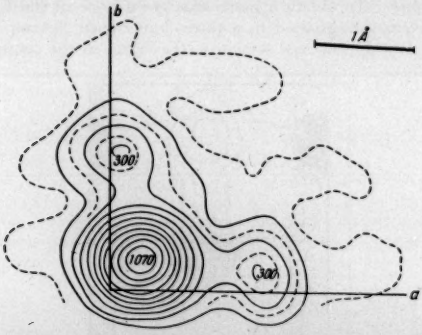


Figure 8. Three-dimensional Fourier section of the potential distribution through the plane of the CH_2 group in paraffin wax. (*After* Weinstein and Pinsker 1950 b.)

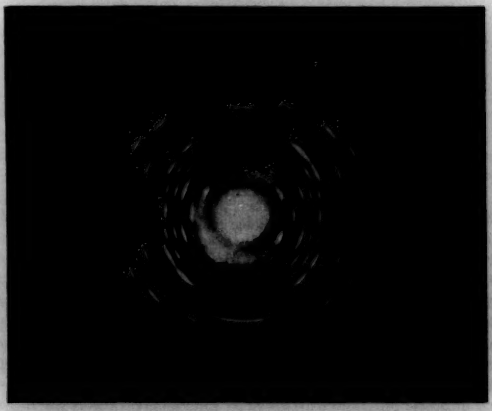


Figure 9. Arc pattern from an inclined and oriented polycrystalline specimen of diketopiperazine. (*After* Weinstein 1955 b.)

the hydrogens accurately, it was necessary to use three-dimensional methods. First a limited projection, or a projection of only that fraction of the unit cell containing the plane of the molecules, was calculated. This showed the hydrogen in the NH–O hydrogen bonds clearly but not the hydrogens in the CH_2 groups.

A complete three-dimensional analysis followed, using the method of generalized projections, and finally sections were calculated through each atomic centre to give the three-dimensional synthesis shown in figure 10. The value of the reliability index, defined by equation (2.19) was 20%. The estimated possible errors in the determination of atomic positions were 0·014 Å for carbon, nitrogen and oxygen and 0·05 Å for hydrogen in projections, while for three-dimensional series the corresponding figures were 0·008 Å and 0·03 Å. The reliability of these estimates may be questioned on the ground that the atom scattering curves are not very accurate and thermal effects have been neglected.

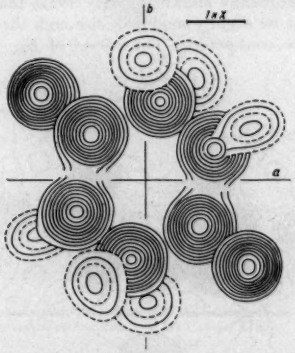


Figure 10. Three-dimensional synthesis of structure of diketopiperazine. (*After* Weinstein 1955 b.)

The interatomic distances found between the heavier atoms were slightly different from those found in the x-ray work of Corey (1938). The C–H distance of 1·09 Å and the N–H distance in the hydrogen bond of 1·00 Å were in agreement with accepted values.

4.4. *Single-crystal Structure Analyses*

The first single-crystal structure analysis was made in connection with an analysis of secondary scattering effects in dicetyl patterns (Cowley, Rees and Spink 1951 a). Since then the method has been developed by the Melbourne group, who have carried out structure analyses from single microcrystals of a number of substances including boric acid, new phases found in oxidized aluminium and gold films, a graphite–ferric chloride compound and basic lead carbonate.

All crystals studied proved to be disordered to a greater or lesser extent. The boric acid and basic lead carbonate crystals, in particular, formed excellent examples of disordered layer-lattice structures with high density of stacking faults. The distribution function technique of analysis, mentioned in § 3 above, was developed in connection with the study of boric acid (Cowley 1953 c) and used again in a different fashion for the structure analysis of basic lead carbonate (Cowley 1956 a).

The normal ordered structure of boric acid, as determined by Zachariasen (1934, 1954) is triclinic, being made up of individual layers of hexagonal symmetry stacked with lateral translations not simply related to the layer unit cell dimensions. The electron diffraction patterns had very nearly hexagonal symmetry (figure 4). This can be expected if the layers are stacked according to a distribution function which has hexagonal or radial symmetry, as reference to (3.13) and (3.14) shows. The two-dimensional form of (3.14)

$$E_{hk} = C_{hk} E_{hk}^{0} \qquad \ldots\ldots (4.8)$$

was used. It was subsequently shown (Cowley 1957) that the interpretation of the C_{hk} and $D(xy)$ must be slightly modified, although the original interpretation is justifiable in the case considered. The values of E_{hk} were determined from

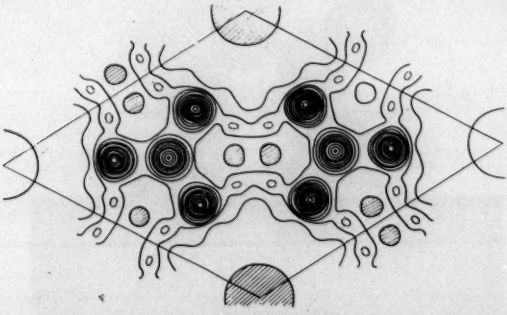


Figure 11. Fourier projection of the potential distribution in a single layer of a disordered boric acid crystal. (*After* Cowley 1953 c.)

the observed intensities. Approximate values of E_{hk}^{0} could be calculated from the x-ray diffraction results, and these were used to calculate the C_{hk} from (4.8). Summing the Fourier series with C_{hk}, assumed to be all positive, as coefficients gave a first approximation to the distribution function $D(xy)$. This was interpreted in terms of sharp peaks, representing preferred inter-layer displacements plus a smooth background due to arbitrary displacements, and fortuitous background fluctuations. Ignoring the background fluctuations an 'idealized' $D(xy)$ was proposed and the corresponding C_{hk} calculated. These were then substituted in (4.8) to give an improved set of E_{hk}^{0}. Continuation of this process led to a refinement of both the layer projection $\phi^{0}(xy)$ (figure 11) and the distribution function projection $D(xy)$ (figure 12). The hydrogen atoms in the O–H ... O hydrogen bonds are clearly revealed in the layer potential projection, but because of the disorder of the structure the hydrogen atom does not appear nearer to one oxygen atom than the other. Instead half a hydrogen atom appears at each of the two possible hydrogen positions 1·00–1·05 Å away from each of the two oxygen atoms. An unexpected feature of the structure is that the O–H ... O hydrogen bonds do not appear to be linear. The hydrogen bonds occur in pairs which

appear to be bowed together as if the hydrogen position is determined not only by the two oxygens of its own hydrogen bond but also by the two oxygens of the neighbouring hydrogen bond.

In the study of basic lead carbonate, $2PbCO_3.Pb(OH)_2$, the distribution function method was used in a more generally justifiable form (Cowley 1956 a). The two-dimensional equivalent of equation (3.15), namely, $I_{hk} = C_{hk}^2 I_{hk}^0$, was used. The C_{hk}^2 are the values at the points hk of the Fourier transform of $Q(xy)$, the projection of the Patterson function corresponding to $D(xy)$. The probable form

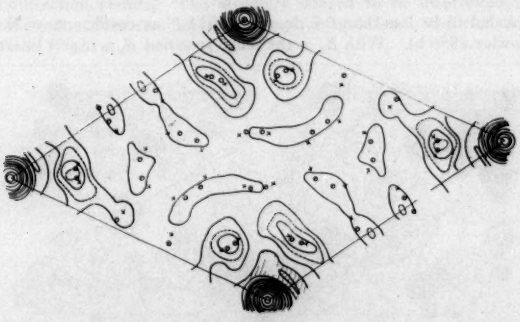


Figure 12. Distribution function projection for disordered boric acid structure at final stage of refinement. (*After* Cowley 1953 c.)

of $Q(xy)$ could be seen from the Patterson function for the whole structure $P(xyz)$ to correspond approximately to that of a rhombohedral stacking of the individual layers. From this C_{hk}^2 values were calculated and a first approximation to the I_{hk}^0 and so $P^0(xy)$ obtained, leading in turn, when interpreted by image-seeking methods, to an approximate layer structure. The structure was then refined by a succession of stages, calculating alternately distribution Patterson functions and projections of the individual layers. The unit cell of the individual layers of the basic lead carbonate structure proved to have trigonal symmetry with $a = b = 9{\cdot}06$ Å, $c = 8{\cdot}27$ Å. A rhombohedral stacking of such layers would give a unit cell with a c axis of 24·8 Å. The Patterson projection $P(xy)$ did not, therefore, contain enough information to assist very much with the structure analysis. Some three-dimensional data were needed. A second projection of the structure along a principal axis of the unit cell was not practicable since it would have involved obtaining patterns with the electron beam parallel to the faces of the thin plate-like crystals used. Hence a ' stereoscopic ' method of analysis (Cowley 1956 c) was employed using Patterson projections calculated from two patterns, one with the beam perpendicular to the crystal plate and one with the beam at an angle of $17\frac{1}{2}°$ to this. The z-coordinate of any Patterson peak was then given by the displacement of the peak in the second projection relative to its position in the first. Sections of the Patterson function for various values of z were calculated by displacing the origin of the tilted Patterson projection by an appropriate

amount and then finding the minimum value from the two projections, taking into account the symmetry of the Patterson function which was, in this case, a six-fold symmetry about the c axis. The two Patterson projections, and two sections of the Patterson function calculated by this stereoscopic method are shown in figure 13 (a), (b), (c), (d).

Actually, instead of Patterson projections as usually defined, modified Patterson projections were used in this work, calculated by using

$$\mathscr{E}_h = |E_h|^2 - 3|E_h|^4 + 4|E_h|^6 - |E_h|^8$$

with E_h scaled to be less than 0·7, in place of $|E_h|^2$ as coefficients of the Fourier series (Cowley 1956 b). With E_h in the range specified $\mathscr{E}_h$ is nearly linear with E_h.

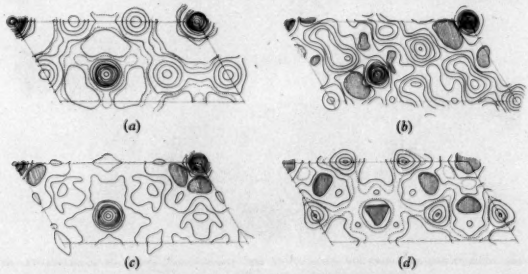


Figure 13. (a, b) Patterson projections of basic lead carbonate: (a) down the c axis; (b) in a direction at 17½° to the c axis. (c, d) Sections of the three-dimensional Patterson function of basic lead carbonate obtained by the stereoscopic method: (c) $z = 0$; (d) $z = 0{\cdot}13$. Hatched areas represent negative values of the functions. (*After* Cowley 1956 c.)

The modified Patterson function is given by adding to the normal Patterson function a series of multiple Patterson functions. It can be shown that the effect is to sharpen up the peaks of the Patterson map and to emphasize the finer detail of the map at the expense of the strongest peaks. Such modified Patterson maps are particularly suited to use with image-seeking techniques.

Once the Patterson projection for a single layer $P^0(xy)$ was obtained, the positions of the lead atoms in the layer unit cell could be deduced without great difficulty by image-seeking and trial-and-error methods. The lead atoms dominated the phases of the structure factors $E_{hk}{}^0$. Fourier projections $\phi^0(xy)$ then showed clearly the positions of the carbon and oxygen atoms. A difference Fourier projection, from which contributions of lead atoms have been subtracted, is shown in figure 14. The individual layers were shown to have the form of a sandwich, with sub-layers of $PbCO_3$ on either side of a hydroxide sub-layer. The final distribution function showed that the stacking of layers deviated appreciably from a rhombohedral stacking in such a way that a fully ordered structure would be triclinic, space group P1. In the crystals examined, however, the deviations from rhombohedral stacking were randomly distributed in such a way that no three-layer unit cell could be defined except on a statistical basis.

As in the case of basic lead carbonate, the study of the ferric chloride–graphite compound by Cowley and Ibers (1956) represented a detailed structure analysis of a substance for which x-ray diffraction studies had yielded no more than diffuse ring patterns containing only a small amount of structural information, which had, in fact, not been interpreted correctly. X-ray powder patterns were actually used in conjunction with the single-crystal electron diffraction data in the study of ferric chloride–graphite, but, except in the determination of inter-layer spacings and layer stacking sequences, served mostly as confirmatory evidence for the electron diffraction results. The structure proved to be disordered in a most unusual manner. Alternate layers of graphite and ferric chloride are stacked

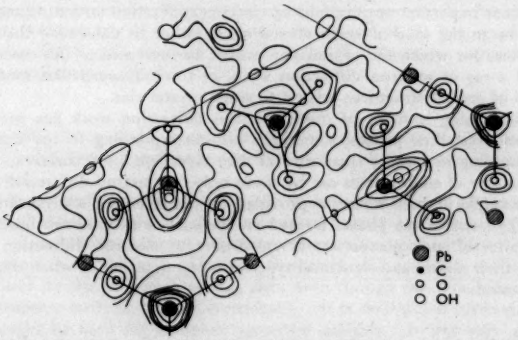


Figure 14. Difference Fourier projection of basic lead carbonate layer. Contributions of all Pb atoms have been removed. (*After* Cowley 1956 a.)

regularly to give a c axis of 9·41 Å. The individual layers of graphite and ferric chloride have hexagonal structure, but the a axes are incomensurate, so that there is no repeat distance in the plane of the layers common to the two-layer types. A ferric chloride layer and a neighbouring graphite layer are related only in that the a axes of the two layers are always at 30° to one another, and local distortions of the ferric chloride configuration take place to allow some correlation between neighbouring chlorine and carbon atom positions. There is some small amount of short range order between the lateral position of one ferric chloride layer and the nearest other ferric chloride layers.

4.5. *Applications of Structure Analysis from Intensities*

The examples of structure analysis quoted above serve to indicate some of the principal fields of application of structure analysis by the methods described. In general, it may be said that the experimental techniques are not so highly developed and intensity measurements are usually not so accurate as in x-ray diffraction work. If, therefore, the desired information may be obtained by the

standard procedures of x-ray diffraction structure analysis, there is no advantage to be gained by using electron diffraction techniques. It has been amply demonstrated that electron diffraction is much more powerful than x-ray diffraction in the detection of light atoms in the presence of heavy atoms, as when hydrogen atoms are to be located in organic compounds, or when light atoms, such as carbon, oxygen and nitrogen, are found in compounds containing lead or platinum. The recent development of neutron diffraction structure analysis techniques has provided an even more powerful technique for these purposes which can be applied when suitable large crystals of the relevant substance are available, but the advantage of electron diffraction work remains when such large crystals cannot be obtained or when suitable neutron sources are not available.

The most important applications of electron diffraction structure analysis will probably be in the field of small crystal sizes, either to determine the structure of substances for which the crystal size cannot be increased to that necessary for successful x-ray or neutron diffraction work, or to study particular modifications or defects of crystal structures related to small crystal size.

In a surprising number of cases, electron diffraction work has revealed the presence of crystalline phases, sometimes not corresponding to any known bulk phase, occurring in minute quantities in thin layers on solid surfaces. A study of the structure of such phases can be made only by electron diffraction methods. A good example of such a case is provided by the structure analysis by Moodie (1957 a, b) of one of the phases formed on the surface of gold foil heated in air. The existence of such phases was revealed only by electron diffraction observations, and their nature and structural type were first determined when this analysis was performed.

§ 5. Structure Analysis from Dynamic Scattering Effects

5.1. *Theory of Dynamic Effects*

It has been convenient and usually adequate to interpret most observations in electron diffraction in terms of the kinematic approximation, but the occurrence of results finding no ready interpretation has stimulated re-examination in terms of the dynamic approximation in which the effect of the crystal potential on the diffracted beams is taken into account. The general theory of dynamic interaction has been given in § 3.2 ; only those aspects relevant to the use of dynamic effects in structure analysis will be discussed here. Dynamic interaction has the general result that the scattering amplitude distributions around the integer points of the reciprocal lattice are no longer independent of the integer point coordinates and the direction of the electron beam in the crystal. The influence of the crystal structure is contained in the Fourier coefficients $E(hkl)$ of the potential. It is this fact that provides the possibility of deducing the structure of crystals from data other than the reflection intensities. MacGillavry (1940) first realized that structure amplitudes could be deduced from such effects. Convergent-beam diffraction patterns obtained from relatively thick (of the order of 900 Å) mica sheets by Kossel and Mollenstedt (1939) contained fine structure in the form of parallel bright and dark bands. The appearance of these fringes suggested an origin in the subsidiary maxima of the kinematic shape transform ; indeed, the

qualitative features of these patterns were interpreted on this basis by Kossel and Mollenstedt. MacGillavry, however, recognized that the positions of the maxima and minima were not in accord with the predictions of kinematic theory and applied Bethe's dynamic theory to the case of a thin parallel-sided crystal of infinite extent. The particular feature of the convergent-beam patterns that could be used to test this theory was the positions of the maxima and minima of density in the banded regions. MacGillavry showed that the minimum densities occurred at angles θ_m, given by the relation

$$\theta_m - \theta_0 = \pm d_{hkl}\left(\frac{m^2}{H^2} - \frac{v_{hkl}^2}{4\pi^2 k}\right)^{1/2}, \qquad \ldots\ldots(5.1)$$

where θ_0 is the angle at which the principal maximum occurs, d_{hkl} the interplanar spacing for (hkl) planes, H the thickness of the crystal, k the propagation factor, $v_{hkl} = (8\pi^2 me/h^2)(E_{hkl}/V)$ and $m = 1, 2, 3$, etc. The kinematic approximation is equivalent to placing $v_{hkl}^2/4\pi^2 k = 0$, i.e.

$$\theta_m - \theta_0 = d_{hkl} m^*/H.$$

It may be seen that a simple geometrical measurement of $(\theta_m - \theta_0)$ allows deduction of v_{hkl} and hence the Fourier coefficient E_{hkl} of the potential distribution in the crystal.

MacGillavry deduced values of E_{hkl} from convergent-beam patterns of mica which agreed satisfactorily with those calculated from x-ray data using Poisson's equation. Considering the approximations necessary in deriving these expressions, the agreement is surprising. However, this work did establish the basis of a new approach to the collection of E_{hkl} data even though the possibility of using this for structure analysis was not considered. It is doubtful if convergent-beam patterns could be used for complete structure analysis. It was only after fine-structural features of diffraction patterns of crystals of fairly regular habit had been discovered and interpreted that structure analysis based on geometrical features of a diffraction pattern became a possibility.

Heidenreich (1949) reformulated the dynamical theory of an infinite parallel-sided crystal in terms of the zone structure of the crystal and applied the theory to the interpretation of extinction contours in electron micrographs of thin crystal sections. He further obtained expressions for the variation of intensity of diffracted beams with distance of propagation through a crystal and obtained a quantitative interpretation of the electron microscope images of thin wedge crystals.

The extension of Bethe's theory to finite polyhedral crystals was made by Kato and Uyeda (1951 a, b) and Kato (1952 a, b) (see also Molière and Niehrs 1954). It is this work which forms the basis of interpretation of the dynamical effects observed in high resolution electron diffraction patterns. As a first step Kato extended Bethe's treatment of the Laue case of the parallel-sided slab to a wedge-shaped crystal of infinite extent and so obtained expressions for the dispersion surface and the wave functions of the transmitted waves. For electron diffraction purposes it is usually more convenient to use Ewald's construction to determine the wave vectors, so that it is necessary to translate the results in the neighbourhood of the Laue point to the appropriate reciprocal lattice point. The relationship between the construction at the dispersion surface and the Ewald construction is

shown in figure 15, following Kato (1952 b). The terminal point of the displaced vector moves along a hyperbolic curve when the crystal orientation deviates from the Bragg angle at a fixed azimuthal angle around the reciprocal lattice vector. The pair of hyperbola so obtained are asymptotic to a pair of intersecting straight lines whose point of intersection is displaced from the reciprocal lattice point.

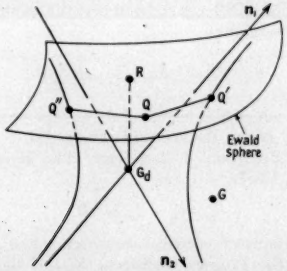


Figure 15. Wave vector relations for dynamic theory represented according to Ewald's construction : $\mathbf{n}_1$, normal to entrance face of crystal ; $\mathbf{n}_2$, normal to exit face of crystal ; G, reciprocal lattice point ; G_d, intersection of vectors representing face normals displaced from G by virtue of mean inner potential of crystal ; Q, end point of vacuum diffracted wave vector if dynamic interaction is neglected and if G is extended along vector direction only ; Q′, Q″, end points of vacuum doubly-diffracted wave vectors after leaving the crystal ; G_dR, normal to Ewald sphere. (*After* Kato 1952 b.)

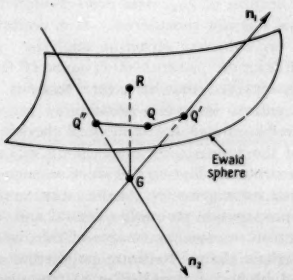


Figure 16. Wave vector relations for kinematic theory represented according to Ewald's construction : $\mathbf{n}_1$, $\mathbf{n}_2$ and G have the same significance as in figure 15 ; GR is the normal to Ewald sphere ; Q is the end point of the vacuum diffracted wave vector if G is extended along vector direction only ; Q′, Q″ are the end points of the vacuum diffracted wave vectors after leaving the crystal. (*After* Kato 1952 b.)

Physically, this indicates (*a*) a simple refraction given by the reciprocal lattice point displacement and (*b*) a double refraction given by the splitting and further displacement from the displaced point. The relation to the kinematic situation may be seen by comparing figure 15 with figure 16, which shows the analogous construction for kinematic theory.

The next step in approaching a physically significant model for a polyhedral crystal was provided by Kato by including in the simple wedge treatment an added boundary condition in the form of limitation of the extent of the entrance and exit faces by finite apertures. These apertures introduce aperture diffraction functions into the wave functions for the waves leaving the crystal. It is equivalent to limitation of the crystal dimensions in directions normal to the incident beams; for small apertures (less than 10 Å) interaction between adjacent regions

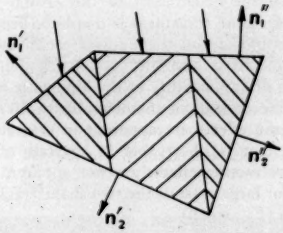


Figure 17. Schematic representation of polyhedral crystal divided into volumes appropriate to different pairs of entrance and exit faces. (*After* Kato and Uyeda 1951 a.)

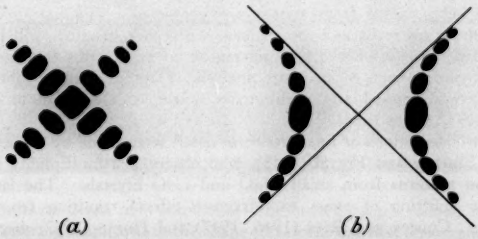


Figure 18. Schematic representation of: (*a*) kinematic shape transform (*after* Rees 1953); (*b*) effective dynamic 'shape transform' (*after* Cowley, Goodman and Rees 1957); for crystal of square cross section.

becomes significant. For larger apertures the aperture diffraction functions become δ-functions and the formulae for the wave functions reduce to those for the infinite wedge. For polyhedral crystals the crystal may be divided into volumes appropriate to pairs of entrance and exit surfaces in figure 17 (Kato and Uyeda 1951 a) and the diffraction in each region considered separately.

At this point it will be profitable to examine the distribution of scattering power in the region of a reciprocal lattice point as it arises in the Ewald construction. The square of the kinematic shape transform describes this for diffraction in the absence of dynamic interaction (see equation (2.9)). For a crystal of square cross section the transform has the form shown schematically in figure 18(*a*)

(Rees 1953), where the black areas represent the regions of non-negligible scattering amplitude. The analogous function in dynamic theory has been termed the 'dynamic shape transform' and for a crystal of square cross section is depicted schematically in figure 18(*b*) (Cowley, Goodman and Rees 1957). The magnitude of the transform is significant along the two branches of the hyperbola determined by the theory quoted earlier. It is important to recognize (i) that the kinematic transform is centred on the reciprocal lattice integer point and is the same for all integer points irrespective of the diffraction conditions, namely, the electron wavelength and the orientation of the crystal relative to the electron beam and (ii) that the dynamic transform is displaced from the integer point and depends in form and amplitude on the Fourier coefficients of the potential E_{hkl}, the wavelength of the electrons and the crystal orientation. Kato (1952 b) has examined the conditions of applicability of the kinematic and dynamic approximations. He found that for crystals of diameter up to 30 Å the theories are indistinguishable, that dynamic effects are apparent in crystals of 50 Å diameter, that the simple refraction effect is observable in crystals of diameter greater than 100 Å and that double refraction effects are not apparent in crystals of diameter less than 270 Å. Even for large crystals the two shape transforms are indistinguishable for weak reflections (E_{hkl} small).

5.2. *Deduction of E_{hkl} from Spot-group Fine Structure*

Kato's theory provides the basis of interpretation of the spot-group fine structure observed in a number of diffraction patterns and hence the possibility of obtaining E_{hkl} values from spot-group geometry. The nature of the fine-structural effects observed and the analysis of the fine structure will be described in a little detail because there is some reason to expect this to develop into an unconventional approach to structure analysis. The advantages inherent in this approach were discussed and preliminary results of Goodman on MgO were reported in 1952 (Rees 1953).

The first observations of fine structural detail were made by Hillier and Baker (1945) and Sturkey and Frevel (1945), who observed a multiplicity in the rings of diffraction patterns from small MgO and CdO crystals. The latter authors ascribed the splitting of rings to refraction effects resulting from the inner potential E_0. Cowley and Rees (1946, 1947) and Honjo (1947) discovered that the individual single-crystal reflections from somewhat larger crystals were in the form of groups of spots disposed about the Bragg point and developed a kinematic theory of refraction to explain their occurrence. Cowley and Rees (1946, 1947) also demonstrated that some of the spot groups exhibited twice as many refraction components as could be accounted for by the simple theory. Examples of this spot-group fine structure are given in figures 19 and 20. Later, this double refraction effect was attributed by Sturkey (1948) to dynamic interaction. A crystal oriented to give the *hkl* reflection would be expected to give, to a reasonable approximation, two effective values of the inner potential, $E_0 \pm E_{hkl}$.

Occasionally, spot groups containing more than two components for each pair of crystal faces may be observed (Cowley and Rees 1946, 1947, Honjo and Mihama 1954). These higher multiplicities possibly originate from higher-order dynamic interactions which are neglected in the two-beam approximation. Honjo and

Mihama (1954) further discuss the relative intensity of components of the doublet corresponding to one pair of crystal faces and attribute the weakening of the outer component to absorption in the crystal, i.e. the potential is to be considered as a complex quantity

$$\phi = \phi_{\text{real}} + i\phi_{\text{imag}}.$$

This fact has not been proposed as a useful contribution to structure analysis, but it will need careful consideration as the technique is evolved.

Figure 19. Spot group from 220 reflection of small single crystal of MgO showing changes in configuration as Ewald sphere is swept through region of reciprocal lattice point. (*After* Cowley, Goodman and Rees 1957.)

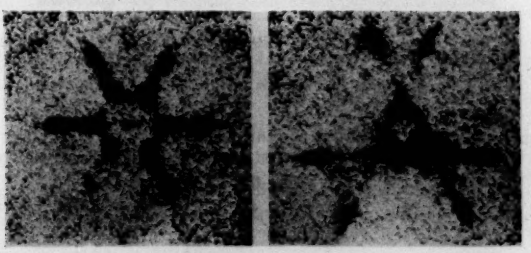


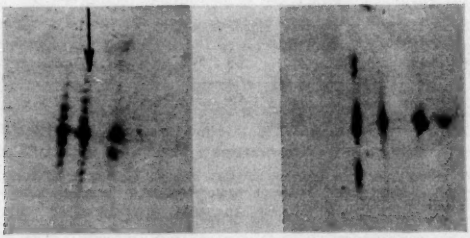
Figure 20. Spot-group fine structure from $\bar{2}20$ reflection from MgO. (*After* Altenhein and Molière 1954.)

Fine-structural detail of a somewhat different character related to the distribution of scattering amplitude in the shape transform was first observed by Rees and Spink (1950 a) in patterns from small ZnO crystals having an elongated hexagonal prismatic habit. Some of these individual reflections consisting of a central maximum with a series of secondary maxima on either side (figure 21(*a*)) could be interpreted satisfactorily in terms of the kinematic shape transform, others (figure 21(*b*)) showed structure suggesting concurrent refraction, a situation again requiring the introduction of dynamic terms.

These fine-structural observations, which have been multiplied by other workers, are interpreted quantitatively in most respects by Kato's dynamical theory. Some residual anomalies, which will be discussed later, probably originate in the approximations which Kato's first-order dynamical theory still contains.

Ideally, structure analysis by this method requires a single crystal of regular habit from which the diffracted beams can be recorded for specific orientations of the crystal relative to the incident beam. The manipulation of crystals of

size 500–1000 Å has not yet been realized, so that it is usually necessary to establish the orientation of a particular crystal from the geometry of the spot group recorded in a polycrystalline pattern. Altenhein, Molière and Niehrs (1952), Molière and Niehrs (1955), Honjo and Mihama (1954) and Cowley, Goodman and Rees (1957) have each proposed methods of deducing crystal orientation and hence E_{hkl} values from spot-group geometry of single-crystal reflections from small cube-shaped crystals.



(*a*) (*b*)

Figure 21. Fine structure of single-crystal reflections from ZnO, showing two examples of transform subsidiary maxima. (*After* Rees and Spink 1950 a.)

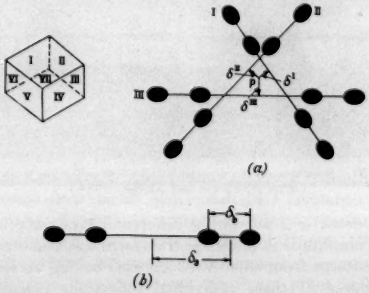


Figure 22. (*a*) Division of cubic crystal in arbitrary orientation into separate diffracting regions and the correlation of spot-group fine structure with this sub-division. (*b*) Definition of parameters characterizing spot separations. (*After* Cowley, Goodman and Rees 1957.)

A crystal of cubic habit in an arbitrary orientation may be divided into regions as in figure 22(*a*). Pairs of faces, such as regions I and IV, contribute four spots along a straight line, so that the spot group from a cubic crystal would be expected to be comprised of three lines of four spots each. One would also expect an undeflected spot from region VII to appear, but this is rarely observed because crystals large enough to give clearly resolved doublets are too thick in region VII to transmit a diffracted beam with any intensity. The diffraction from other regions is limited to thin strips along the crystal edges.

The four spots in each line have separations characterized by two parameters δ_a and δ_b as shown in figure 22(*b*). The vector displacement $\boldsymbol{\delta}_a$ is that corresponding to simple kinematic refraction and is given by (Cowley and Rees 1946, 1947)

$$\boldsymbol{\delta}_a = \frac{E_\theta}{2W}\left(\frac{\mathbf{n}_1}{\cos\chi_1} + \frac{\mathbf{n}_2}{\cos\chi_2}\right), \qquad \ldots\ldots(5.2)$$

where $\mathbf{n}_1$ and $\mathbf{n}_2$ are the face normal vectors and χ_1 and χ_2 are the angles made by the beam with these normals.

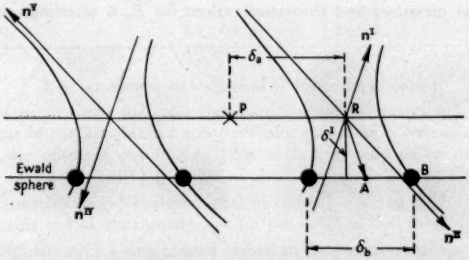


Figure 23. Intersection of the Ewald sphere with effective dynamic shape transform and the relationship to spot separation parameters. (*After* Cowley, Goodman and Rees 1957.)

The vectors $\boldsymbol{\delta}_b'$ and $\boldsymbol{\delta}_b''$ are the dynamic displacements from the kinematic refraction spot (see figure 23) and may be shown to be (Kato 1949)

$$\boldsymbol{\delta}_b'(\boldsymbol{\delta}_b'') = \frac{E_{hkl}}{2W}\left\{(\Lambda^2+1)^{1/2}\left(\frac{\mathbf{n}_1}{\cos\chi_1} + \frac{\mathbf{n}_2}{\cos\chi_2}\right) \mp \left(\frac{\mathbf{n}_1}{\cos\chi_1} - \frac{\mathbf{n}_2}{\cos\chi_2}\right)\right\} \qquad \ldots\ldots(5.3)$$

where

$$\Lambda = -\frac{2\pi}{\lambda}\frac{h^2}{8\pi^2 me}\frac{\zeta}{E_{hkl}}$$

and ζ is the excitation error. The hyperbolic and linear terms of $\boldsymbol{\delta}_b'$ in equation (5.3) are AB and RA respectively in figure 23.

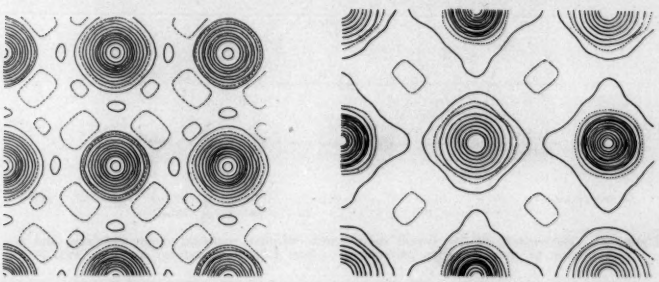
The separation

$$\delta_b = |\boldsymbol{\delta}_b' - \boldsymbol{\delta}_b''| = \frac{2E_{hkl}}{E_0}\delta_a(\Lambda^2+1)^{1/2}. \qquad \ldots\ldots(5.4)$$

δ_b/δ_a is a hyperbolic function of Λ with asymptotes $\pm\Lambda 2E_{hkl}/E_0$. At the Bragg angle $\Lambda = 0$ and $\delta_b/\delta_a = 2E_{hkl}/E_0$, which is Sturkey's (1948) result.

Essentially similar methods of deriving the E_{hkl} and E_0 values from individual spot groups were used by Honjo and Mihama (1954), Altenhein, Molière and Niehrs (1952), Molière and Niehrs (1954) and Cowley, Goodman and Rees (1957). The latter authors adopted the experimental technique of rotating the specimen through incremental angles of the order of 10^{-3} radian between exposures. This allows a series of observations to be made on each spot group. A selection from one such series illustrating the change in configuration of the spot group is given in figure 19. Values of χ_1, χ_2, $\mathbf{n}_1$, $\mathbf{n}_2$, and Λ can all be calculated from the

geometry of the configuration except in the case of $h00$ reflections where an E_0 value must be assumed. From a plot of δ_b against Λ, which, from equation (5.4), is expected to be hyperbolic, it is possible to interpolate a value of δ_b for $\Lambda = 0$. From δ_a and δ_b ($\Lambda = 0$) it is then possible to compute E_0 and E_{hkl}. Experimental and theoretical values (computed from known structure) for MgO are given in the table. The theoretical values have been corrected by a temperature factor of the conventional form. In view of the likely experimental errors and the approximate nature of the theory used in interpreting spot groups, the correspondence between measured and theoretical values is satisfactory. The consistent discrepancy between measured and theoretical values for E_0 is puzzling, since there is



(*a*) (*b*)

Figure 24. Sections at $z = 0$ of potential distribution in MgO using structure amplitudes obtained from spot-group fine structure. (*a*) Experimental E_{hkl} data. (*b*) Experimental E_{hkl} with estimated E_{hkl} values for weak reflections added. (*After* Cowley, Goodman and Rees 1957.)

no reason to suspect errors of the order of 2–3 ev in the theoretical value. The two-beam dynamic approximation is one obvious cause of error; moreover, the way in which the two-beam approximation is applied is in itself approximate. Again there is a progressive divergence of theoretical and experimental E_{hkl} values with increase of $\sin\theta/\lambda$, corresponding to a much higher effective temperature factor than that used in the calculation of the theoretical values. Investigation of the source of these discrepancies must be the first step in any move to develop this approach to structure analysis.

To establish the possibilities of structure analysis by this method, Cowley, Goodman and Rees (1957) calculated Fourier sections of the potential distribution for the $z = 0$ plane of the MgO unit cell. The summation

$$\sum_h \sum_k \sum_l E_{hkl} \cos 2\pi(hx + ky)$$

did not include any of the weak h, k, l odd terms other than (111), because this technique at present is not adequate for weak reflections. Such a summation is represented in the contour map of figure 24(*a*). This may be compared with the section of the distribution obtained when theoretical E_{hkl} values for h, k, l odd are included (figure 24(*b*)).

Experimental and Theoretical Values of E_{hkl} for MgO (in Volts)

hkl	Honjo and Mihama (1954)	Molière and Niehrs (1955)	Cowley, Goodman and Rees (1957)	Theoretical
000	15·5	15·3 ± 0·2	14·4 ± 0·3	12
111	1·7	1·25 ± 0·15	1·2 ± 0·2	1·29
200	7·0	7·55 ± 0·2	8·0 ± 0·4	7·58
220	5·5	5·09 ± 0·13	5·3 ± 0·2	5·02
222	4·2	3·5 ± 0·3	3·5 ± 0·2	3·80
400	—	2·6 ± 0·6	2·7 ± 0·4	3·09
420	—	2·4 ± 0·3	2·2 ± 0·15	2·57
422	—	2·0 ± 0·7	1·5 ± 0·2	2·20
440	—	2·1 ± 0·4	1·2 ± 0·2	1·65
442	—	2·0 ± 0·6	—	1·48

5.3. *Assessment as a Method of Structure Analysis*

The present situation indicates that partial structure analysis is possible and an assessment of the possibilities and the major problems is warranted.

Ideally, the practical requirements for such a method to be useful are the following:

(i) The substance to be analysed must be capable of being produced in crystals of regular habit and of appropriate size (of the order of 500 to 1000 Å diameter).

(ii) Manipulation of a single small crystal under electron microscopical observation to set it in an optimum orientation should be possible.

(iii) The electron diffraction camera should have sufficient resolving power to provide E_{hkl} values for very weak reflections.

Requirement (i) is satisfied beautifully by MgO and CdO, but it is not easy to induce many substances to adopt a regular habit and this aspect will require considerable investigation. However, one is not required, as in x-ray diffraction, to produce crystals of about 0·5 mm in diameter, but about a hundred times smaller.

It is not possible to meet requirement (ii) at present, both because sufficiently refined manipulative techniques are not available and because of limitations of electron intensity. The specimen manipulator should not be an insuperable instrumentation problem. However, the task of providing greater electron intensity for this and requirement (iii) appears to have no obvious solution. It would be necessary to observe a spot group during the manipulation to allow observation of the optimum spot-group configuration. Moreover, if E_{hkl} values smaller by a factor s than those capable of measurement at present are to be recorded, then the smaller effective source size would require an increase in exposure time by s^4. This is quite impracticable even for an improvement by a factor 2. Higher brightness electron sources or much more sensitive detectors are the only possibilities.

The principal advantages of this approach to structure analysis are (i) that the analysis involves measurement of angles and distances rather than intensities, (ii) that extremely small crystals may be studied, and (iii) that, provided the habit is regular, the shape is not restricted. In this latter respect this method is to be compared with the orthodox approach, which is substantially limited to crystals

of layer–lattice structure crystallizing in extensive thin plates. It is further noteworthy that polycrystalline specimens may be used as in the present series of investigations. Because of the several factors influencing the distribution of intensity across electron diffraction rings (Rees and Spink 1950 a), it is extremely difficult to obtain reliable relative intensities from continuous ring patterns.

The difficulties confronting this unorthodox approach to structure analysis are technical and it is not unreasonable to expect these to be overcome in the near future and so provide a powerful new method.

§ 6. Direct Observation of Crystal Lattice Structure

The direct observation of the structure of a crystal lattice by means of fast electrons has for long been an unrealized but theoretically possible goal. The last few years have seen three independent lines of work open up this possibility. None of these approaches has yielded more than an indication of the possibilities, but the physical basis for each is well attested on theoretical and experimental grounds. It is certain that the next few years will see the technique of these methods developed to a point where at least one of them will yield direct and complete resolution of the atomic arrangement in crystals. The three methods are :

(i) Direct imaging of crystals containing heavy atoms in the electron microscope, realized in part first by Menter (1956 a).

(ii) Making use of a new imaging phenomenon for periodic objects, termed Fourier imaging and first described and discussed theoretically by Cowley and Moodie (1957 a, b, c).

(iii) Using electron interferences between electron beams passing through successive thin crystals.

At first sight none of these would appear to warrant inclusion in an article on Fourier methods in structure analysis, but it will become evident during the subsequent discussion that the physical processes by which these direct images are formed are very closely related to the Fourier methods used in the computation of structures from diffraction data.

6.1. *Electron Microscope Images of Lattice Structures*

The possibility of resolving isolated atoms in an electron microscope has been the subject of theoretical discussion from time to time. It is now generally agreed that with an electron microscope capable of resolving 3 Å the heavier atoms should be resolvable. With this in mind, Menter (1956 a) examined suitable crystalline substances, namely, the phthalocyanines, selected because they contain characteristic planes of atoms separated by distances (around 10 Å) of the order of the resolution limit of the electron microscope. Crystals of platinum phthalocyanine, having sheets of platinum atoms 11·94 Å apart, and copper phthalocyanine, in which the copper atoms are 9·8 Å apart, gave images crossed by parallel lines uniformly spaced. In the platinum compound these lines were $12{\cdot}0 \pm 0{\cdot}2$ Å apart and in the copper compound $10{\cdot}30 \pm 0{\cdot}3$ Å. The lines are parallel to the long edge of the ribbon crystals, i.e. [010] and the spacing is correlated with the $(20\bar{1})$ planes, although some unexplained discrepancies exist for the copper derivative. These micrographs, of which figure 25 is a typical example,

are recorded with a limiting objective aperture of 50μ, corresponding to a semi-angular aperture of about 10^{-2} radian. Planes of spacing less than 4·2 Å do not then contribute spectra to the image and in fact the only spectra involved are $(20\bar{1}), (\bar{2}01)$, for which $d = 11{\cdot}94$ Å, and $(40\bar{2}), (\bar{4}02)$, for which $d = 5{\cdot}97$ Å. It is the recombination of these spectra and the zero-order spectrum in the image plane that Menter suggests produces the structure in the image. Menter has discussed the theory of this image formation in terms of Abbe's theory. According to this an ideal lens will resolve a spacing d when the first-order diffracted beam from this spacing is passed by the aperture. This limiting condition is

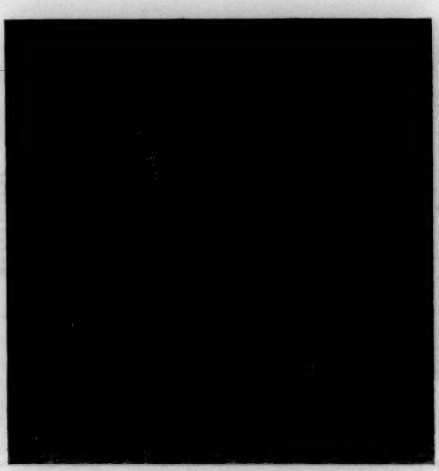


Figure 25. Electron micrograph of platinum phthalocyanine crystal .showing structure of $(20\bar{1})$ planes ($\times 1\,000\,000$). (*After* Menter 1956 a)

$d = \lambda/\alpha$, where α is the angle between the diffracted and incident beams. However, the order of resolution obtained in these micrographs would at first sight appear better than the known spherical aberration of the electron microscopic objective would allow. The diffracted beams pass through a narrow zone of the objective and the phase delay across the wave front is calculated to be negligible. The path difference $\Delta\epsilon$ arising from the spherical aberration C_s across a point Δr wide at a distance r from the axis of a lens of focal length f is $\Delta\epsilon = C_s r^3 \Delta r / f^4$, which is much smaller than λ for the conditions under which Menter obtained his micrographs. If $\Delta\epsilon$ is much smaller than λ, interference between the diffracted and zero-order beams may be expected to produce an image of the diffracting planes. The overall phase delay of the wave front affects the image contrast only. Phase shift originating in defect of focus may contribute to increase of image contrast. Menter's view of his results is that the image is a projection

of the crystal structure in the direction of the electron beam. Since the phase delays for more diffracted beams passing through different zones of the objective will necessarily be different, it is unlikely that a complete image of the crystal structure will be attained. The method has been extremely successful in providing the first direct observation of edge dislocations and other crystal imperfections (figure 26). Menter (1956 b) has recently reported further extension of these results to the imaging of two sets of planes spaced 14·5 Å apart intersecting at 70° in sodium faujasite and resolution of the 6·8 Å planes in their molybdenum trioxide crystals.

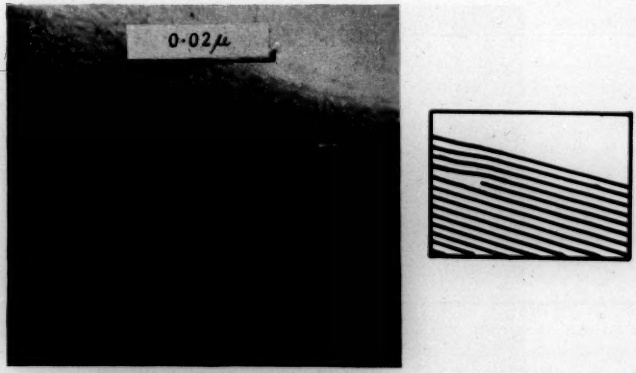


Figure 26. (*a*) Electron micrograph of platinum phthalocyanine crystal showing a single edge dislocation (×1 500 000). (*After* Menter 1956 a.) (*b*) Diagrammatic representation of delineated area in (*a*).

One point overlooked by Menter in his discussion of the formation of these images is that a crystal of thickness of about 100 Å will act predominantly as a phase grating, since any absorption leading to the amplitude-grating function will be negligibly small (Boersch 1948). Menter's lattice planes will then appear in out-of-focus images only; in-focus images will show no structure. Such out-of-focus images may not always represent a projection of the crystal lattice, but a somewhat more complex distribution.

6.2. *Fourier Images*

Cowley and Moodie (1957 a, b, c) have recently examined the possibility of using the *a priori* knowledge of periodicity in a crystal to design an optical system capable of forming an image of the periodic structure of the crystal. It was predicted that such a system was theoretically possible and simple experiments using an optical analogue confirmed these predictions. It is possible to use a periodic

object as its own image-forming system provided certain fairly simple conditions are satisfied.

The simplest model for the purposes of explaining the phenomenon is that of a point source of radiation and an infinite plane grating. The result of this analysis is that there exist planes parallel to the plane of the grating on which images of the grating are reconstituted by interference between the diffracted rays under conditions in which path lengths differ by integral numbers of wavelengths.

If a one-dimensional periodic object is defined by

$$\sum_h F_h \cos\left(h\frac{\zeta}{a}+\phi_h\right)$$

then it is found that the intensity $I(x)$ at a point x in the plane of observation is given by

$$I(x) = K_2^2\left[\sum_h F_h^2\cos^2\psi_h + \sum_i\sum_j \cos(\gamma_i^2-\gamma_j^2)F_iF_j\cos\psi_i\cos\psi_j\right], \quad \ldots\ldots(6.1)$$

where the F's are structure amplitudes,

$$\gamma_h = \pi\lambda\frac{R+R_q}{RR_q}\frac{h^2}{a^2},$$

$$\psi_h = 2\pi\left(\frac{R_q}{R+R_q}\frac{hx}{a}+\phi_h\right),$$

K_2 is a constant which may be left undefined, ϕ_h is the phase angle associated with F_h, ζ, x are coordinates on the object and observational planes respectively, a is the cell parameter, and R and R_q are the normal distances from the object plane to the plane of observation and the source respectively.

It is found that the conditions

$$\cos(\gamma_1^2-\gamma_2^2) = \ldots = \cos(\gamma_i^2-\gamma_j^2) = +1$$

define a set of observational planes for which the intensity is that due to the distribution

$$K_2\sum_h F_h\cos\psi_h,$$

which is that of the original distribution magnified by the factor $(R+R_q)/R_q$. At intermediate planes the intensity distribution bears no simple relation to that of the object.

The above conditions may also be written in the more physically significant form

$$\frac{1}{R}+\frac{1}{R_q} = \frac{1}{n}\frac{\lambda}{2a^2} \qquad \ldots\ldots(6.2)$$

where n is any positive or negative integer. Equation (6.2) is simply that of an image-forming component possessing multiple focal lengths $n2a^2/\lambda$. The images formed on these focal planes have been termed Fourier images. It might be expected that the phenomenon is restricted to periodic objects and that non-periodic regions in such objects will tend to be averaged out.

Cowley and Moodie (1957 b, c) have extended their theoretical examination to two dimensions and have investigated the effect of finite source size, coherence and aberrations in auxiliary optical systems so that the possible application to the imaging of real crystals could be assessed. The limitation of the theoretical treatment to plane gratings is equivalent to the approximation of the Ewald sphere by a plane, which is adequate for fast electrons and real crystals.

The most important result of these investigations is that the diameter of a rectangular source may be many times greater than the smallest distance to be resolved in the crystal, provided that the edge of the source is defined to that order. Cowley and Moodie have considered various ways of realizing sources of the required properties. However, it would not appear to be practicable to obtain the whole magnification in a single Fourier imaging step, since the source would need to be too close to the object for practical realization. If R_q is chosen to be of the order of a micron, then Fourier magnifications large enough to bring lattice periodicities within the resolution capabilities of electron microscopes can be readily attained. Using Fourier imaging for the primary magnification in a conventional electron microscope would appear to represent the best prospect for direct observation of crystal structure.

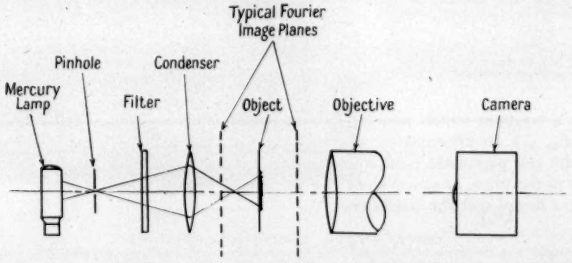


Figure 27. Optical arrangement for observing and recording Fourier images of two-dimensional periodic objects. (*After* Cowley and Moodie 1957 a.)

The least resolvable distance Δ in such a system is given by

$$\Delta = [\Delta_f^2 + (\Delta_m/M)^2]^{1/2}, \qquad \ldots\ldots(6.3)$$

where Δ_f is the least resolvable distance in the Fourier image of magnification M and Δ_m is that for the electron microscope. For a source having an edge unevenness of σ, Δ_f is given by

$$\Delta_f = R\sigma/(R+R_q).$$

Here Δ, Δ_f, Δ_m and σ are all measured in terms of the lattice parameter.

The feasibility of such a method may be appreciated by a simple numerical example. For a Fourier magnification of 10 times and an average source edge unevenness of one lattice spacing $\Delta_f \sim 1$. For high quality electron microscopes one can expect a performance equivalent to resolution of 5 lattice spacings, under which conditions $\Delta = 1$.

The demonstration of the Fourier imaging phenomenon and the confirmation of the theoretical predictions regarding resolution and source size have been provided by the optical analogue. To make convenient comparison between the real image of a two-dimensional amplitude grating and its Fourier images an ordinary microscope was used in such a way that the condenser focused an image of a pinhole immediately behind the object (see figure 27). Racking the microscope objective up then allowed observation of Fourier images in sequence. Figures 28 (*a*, *b*) show the 4th Fourier image and the real image enlarged to the same magnification. The superior resolution and the suppression of the gross defect in periodicity provided by the Fourier image are readily appreciated.

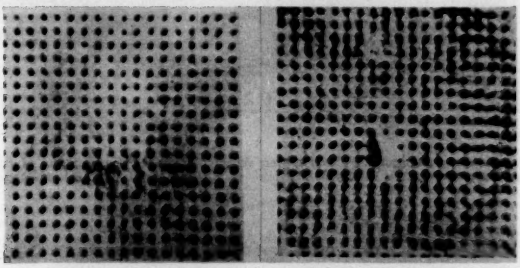


Figure 28. (*a*) The fourth Fourier image of a two-dimensional grating. (*b*) The real image of the same grating enlarged to the same magnification as the Fourier image in (*a*). Note the aperiodic detail in the object and the image defects which are not apparent in (*a*). (*After* Cowley and Moodie 1957 a.)

That the reconstituted image is a true representation of the periodic structure and not the Patterson function has been established theoretically and confirmed by experiments with gratings for which the Patterson function is unequivocally differentiated from the original distribution.

6.3. *Scattering by Successive Crystals*

Observations of fringes across the overlapping regions of thin lamellar crystals in electron micrographs were made in several laboratories during 1950 (figures 29 and 30). These fringes were interpreted as electron interference fringes in the first report by Mitsuichi, Nagasaki and Uyeda (1951), who assumed a large angle of tilt between the two crystals contributing and interference between the transmitted beam and a beam doubly diffracted by two different unrelated reflections in the two crystals. This interpretation did not give very reasonable models of the overlapping crystals; for instance, wedge angles of the order of 5° were necessary to explain the fringe spacings. A more complete examination of the origin of these fringes was performed by Dowell, Farrant and Rees (1954), who demonstrated that rotational disorientation of substantially parallel crystals was a more general cause. The rotational angle ϵ between the two sets of crystals' axes and the fringe spacing s_{hk0} are related by

$$s_{hk0} = d_{hk0}/\epsilon$$

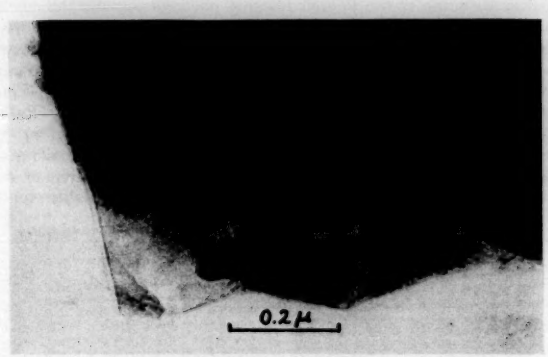


Figure 29. Electron micrograph of overlapping sheets of graphite showing sets of fringes comprising terms of the Patterson distribution. (*After* Dowell, Farrant and Rees.)

Figure 30. Electron micrograph of crossed lamellar crystals of MoO_2 showing overlapping sets of fringes. An edge dislocation is visible near the centre of the field. (*After* Dowell, Farrant and Rees 1954.)

where the interference occurs between the transmitted beam vector $\mathbf{S}_0$ and the doubly diffracted vector $\mathbf{S}_{h,k,0/\bar{h},\bar{k},0}$, the subscripts referring to the two crystals (see figure 31). The relationships are a little more complicated in cases where crystals are crossed (e.g. figure 30) or where two crystals of different structure overlap, but they are readily derived. The significant feature of the above relation is that the fringes are in fact 'images' of the lattice planes ($hk0$) magnified by a factor $1/\epsilon$. If ϵ is small ($<0{\cdot}1$ radian), then images of lattice planes, magnified ten times or more, are well within the resolution capabilities of high-resolution

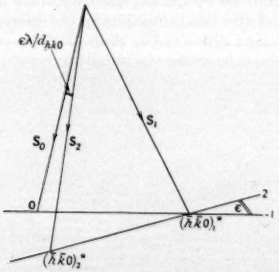


Figure 31. Reciprocal lattice construction for two parallel crystals with rotational disorientation showing angle between incident beam vector and doubly-diffracted beam vector $\mathbf{S}_{hk0/\bar{h}\bar{k}0}(=\mathbf{S}_2)$. (*After* Dowell, Farrant and Rees 1956.)

electron microscopes. Fringes down to 20 Å are frequently observed (Dowell, Farrant and Rees 1954) and recently Bassett, Menter and Pashley (1957) have recorded fringes down to 6 Å. Frequently, images are recorded showing overlapping sets of fringes and this suggested some connection between the complex fringe pattern and the crystal structure. Farrant and Rees (1956) were able to establish theoretically that the pattern was in fact very closely related to the Patterson function of the potential distribution in the crystal and the well-known moiré pattern. The simple theory given here follows Dowell, Farrant and Rees (1956). To a sufficiently good approximation, as discussed earlier, the potential distribution of a thin crystal may be represented by the two-dimensional Fourier series

$$\phi(xy) = \sum_h \sum_k E_{hk0} \exp\left\{2\pi i\left(\frac{hx}{a}+\frac{ky}{b}\right)\right\}.$$

Two identical distributions $\phi(x'y')$ and $\phi(x''y'')$, referred to a common origin, have axes rotated through an angle ϵ with respect to one another. For small ϵ the transformations to a coordinate system with axes forming the bisectrixes between x' and x'' and y' and y'' are given by

$$x' = x+y\epsilon/2\,; \quad y' = -x\epsilon/2+y$$

$$x'' = x-y\epsilon/2\,; \quad y'' = x\epsilon/2+y.$$

Provided the crystals are not separated by more than a few hundred ångström units, then the distribution of amplitude in the image plane is given by

$$\exp\{i\sigma[\phi(x'y')+\phi(x''y'')]\}$$

if the objective lens is perfect and is not restricted in aperture. The expansion of this exponential contains terms representing contributions from the separate crystals and a product term $\phi(x'y')\phi(x''y'')$ representing the interaction of waves scattered by the two crystals. This product is the term of interest. In practice beams travelling in directions at angles greater than 0·005 radian away from the paraxial ray are either removed by the objective aperture or are displaced, through the spherical aberration of the lens, to points in the image plane remote from the image of the crystal. Beams diffracted or doubly-diffracted at angles greater than 0·005 radian are therefore effectively prevented from contributing to the image.

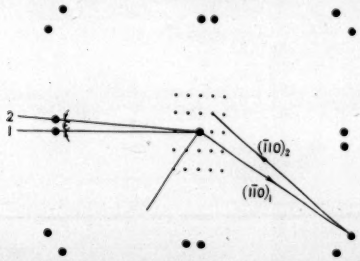


Figure 32. Diffraction pattern of two parallel crystals rotationally disoriented. The primary gratings are shown as heavy points and the miniature secondary grating around the incident beam spot as light points. The origin of a typical secondary grating spot $\{(1\bar{1}0)_1/(\bar{1}10)_2\}$ is illustrated. Secondary gratings will also occur around each of the primary grating points. (*After* Dowell, Farrant and Rees 1956.)

In practice, this means that only those beams corresponding to double diffractions for which $h_2k_2 = \bar{h}_1\bar{k}_1$ contribute to the image. If terms other than these are removed from the product $\phi(x'y')\phi(x''y'')$, then the distribution of amplitude in the image plane for a centro-symmetric crystal is

$$\psi(xy) = \sum_h \sum_k E_{hk}{}^2 \cos\left\{2\pi\epsilon\left(\frac{k}{b}x - \frac{h}{a}y\right)\right\}, \qquad \ldots\ldots(6.4)$$

since $E_{hk} = E_{\bar{h}\bar{k}}$.

Equation (6.4) is simply the Patterson distribution rotated through 90° with respect to the mean axes of the original distributions and having unit translations a/ϵ and b/ϵ. The Patterson distribution is at a magnification $1/\epsilon$. If ϵ is less than 0·1 radian then detail corresponding to vector displacements of the order of 1 Å in the original distribution will be resolved. The diffraction pattern from such a pair of crystals is in the nature of a superlattice pattern and will consist of miniature cross grating patterns of secondary origin around each of the primary reflections (figure 32). The effective aperture of the objective passes only that miniature grating around the incident beam direction, which is of course the

Fourier transform of the distribution $I(xy)$. Diffraction patterns from overlapping crystals giving fringes confirm this.

Analogous treatments for two crystals of the same symmetry in parallel orientation but of different lattice parameters (isomorphous pair) and for two crystals of different structure have been given by Dowell, Farrant and Rees (to be published) and Cowley and Moodie (to be published). If the structure of one crystal of a pair is known, then the convolution function which the moiré image represents may be analysed subsequently by methods analogous to image-seeking. It is seen that this phenomenon represents a promising approach to the direct observation of crystal structure, since it is not limited by the finite size of the electron source or the normal magnitude of the spherical aberration of the objective.

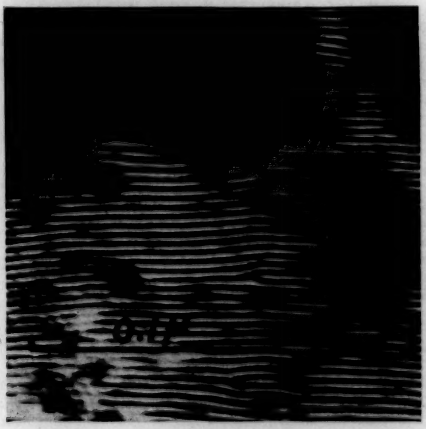


Figure 33. Electron micrograph of specimen composed of thin gold and palladium single-crystal films in parallel orientation showing interference fringes arising through the different dimensions of the two cubic lattices. Several examples of edge dislocations are clearly visible. (*After* Bassett, Menter and Pashley 1957.)

Cowley and Moodie have used their new formulation of electron scattering processes to provide a rigorous theoretical treatment of the origin of the moiré fringes. They have also made an examination of the conditions which must be fulfilled to obtain adequate resolution in the convolution images. These authors and Dowell, Farrant and Rees have examined the theoretical possibilities of using three parallel crystals to produce a convolution image which bears a close relationship to the original potential distribution. The conditions which must be met are of course considerably more severe than for two crystals, but the possible use of such an arrangement is being investigated.

Recently, Hashimoto and Uyeda (1957) and Bassett, Menter and Pashley (1957) have reported on the imaging of dislocations. Both groups of authors

established that a dislocation would be recognizably indicated in the moiré pattern and have interpreted their micrographs along these lines. Bassett, Menter and Pashley obtained beautiful moiré images of edge dislocations (figure 33) by preparing composite film specimens of two metals, e.g. Pd evaporated on to a single-crystal gold film in such a way that oriented growth of the Pd occurs.

The possibility of using two-crystal interference as a direct approach to structure determination has been discussed by Dowell, Farrant and Rees (1956). The technical difficulties of handling crystals or sections of crystals in such a way as to control parallelism and relative orientation does not appear to be an insuperable experimental problem.

References

Altenhein, H. J., Molière, K., and Niehrs, H., 1952, *Naturwissenschaften*, **39**, 447.

Bassett, G. A., Menter, J. W., and Pashley, D. W., 1957, *Nature, Lond.*, **179**, 752.

Bethe, H. A., 1928, *Ann. Phys., Lpz.*, **87**, 55.

Blackman, M., 1939, *Proc. Roy. Soc.* A, **173**, 68; 1951 a, *Proc. Phys. Soc.* B, **64**, 625; 1951 b, *Ibid.*, **64**, 631.

Boersch, H., 1948, *Optik, Stuttgart*, **3**, 24.

Corey, R. B., 1938, *J. Amer. Chem. Soc.*, **70**, 1568.

Cowley, J. M., 1953 a, *Acta Cryst., Camb.*, **6**, 53; 1953 b, *Ibid.*, **6**, 615; 1953 c, *Ibid.*, **6**, 522; 1953 d, *Ibid.*, **6**, 846; 1954, *J. Electrochem. Soc.*, **101**, 277; 1956 a, *Acta Cryst., Camb.*, **9**, 391; 1956 b, *Ibid.*, **9**, 397; 1956 c, *Ibid.*, **9**, 399; 1957, *Ibid.*, **10**, 141.

Cowley, J. M., Goodman, P., and Rees, A. L. G., 1957, *Acta Cryst., Camb.*, **10**, 19.

Cowley, J. M., and Ibers, J. A., 1956, *Acta Cryst., Camb.*, **9**, 421.

Cowley, J. M., and Moodie, A. F., 1957 a, *Proc. Phys. Soc.* B, **70**, 486; 1957 b, *Ibid.*, **70**, 497; 1957 c, *Ibid.*, **70**, 505; 1957 d, *Acta Cryst., Camb.*, **10**, 609; 1958, *Proc. Phys. Soc.*, **71**, 533.

Cowley, J. M., and Rees, A. L. G., 1946, *Nature, Lond.*, **158**, 550; 1947, *Proc. Phys. Soc.* **59**, 283.

Cowley, J. M., Rees, A. L. G., and Spink, J. A., 1951 a, *Proc. Phys. Soc.* A, **64**, 609; 1951 b, *Ibid.* B, **64**, 638.

Dowell, W. C. T., Farrant, J. L., and Rees, A. L. G., 1954, *Proceedings of the 3rd International Conference on Electron Microscopy, London*, 279; 1956, *Proceedings of the Regional Conference on Electron Microscopy, Tokyo*, 320.

Farrant, J. L., and Rees, A. L. G., 1956, *Paper to International Union Crystallography Symposium, Madrid.*

Gabor, D., 1956, *Rev. Mod. Phys.*, **28**, 260.

Gevers, R., 1954, *Acta Cryst., Camb.*, **7**, 740.

Glauber, R., and Schomaker, V., 1953, *Phys. Rev.*, **89**, 667.

Gottsche, H., 1956, *Acta Cryst., Camb.*, **9**, 179.

Hashimoto, H., and Uyeda, R., 1957, *Acta Cryst., Camb.*, **10**, 143.

Heidenreich, R. D., 1949, *J. Appl. Phys.*, **20**, 993; 1950, *Phys. Rev.*, **77**, 271.

Hendricks, S. B., and Teller, E., 1942, *J. Chem. Phys.*, **10**, 147.

Hillier, J., and Baker, R. F., 1945, *Phys. Rev.*, **68**, 98.

Hoerni, J., 1957, *Acta Cryst., Camb.*, **10**, 858 (A).

Hoerni, J., and Ibers, J. A., 1953, *Phys. Rev.*, **91**, 1182.

Honjo, G., 1947, *J. Phys. Soc. Japan*, **2**, 133.

Honjo, G., and Mihama, K., 1954, *J. Phys. Soc. Japan*, **9**, 184.

Honjo, G., and Miyake, S., 1957, *Paper to Electron Diffraction Symposium*, 4th Congress, International Union of Crystallography.

Ibers, J. A., 1957, *Acta Cryst., Camb.*, **10**, 858 (A).

Ibers, J. A., and Hoerni, J., 1954, *Acta Cryst., Camb.*, **7**, 405.

Jagodzinski, H., 1949 a, *Acta Cryst., Camb.*, **2**, 201; 1949 b, *Ibid.*, **2**, 208.

JAMES, R. W., 1948, *The Optical Principles of the Diffraction of X-rays* (London : Bell).
KAKINOKI, J., and KOMURA, Y., 1952, *J. Phys. Soc. Japan*, **7**, 30.
KATO, N., 1949, *Proc. Imp. Acad. Japan*, **25**, 41 ; 1952 a, *J. Phys. Soc. Japan*, **7**, 397 ; 1952 b, *Ibid.*, **7**, 406.
KATO, N., and UYEDA, R., 1951 a, *Acta Cryst., Camb.*, **4**, 227 ; 1951 b, *Ibid.*, **4**, 229.
KOSSEL, W., and MOLLENSTEDT, G., 1939, *Ann. Phys., Lpz.*, **36**, 113.
KUWABARA, S., 1957, *J. Phys. Soc. Japan*, **12**, 637.
LENNANDER, S., 1954, *Ark. Fys.*, **8**, 551.
LIPSON, H., and COCHRANE, W., 1953, *The Determination of Crystal Structures* (London : Bell).
MACGILLAVRY, C. H., 1940, *Physica*, **7**, 329.
MENTER, J. W., 1956 a, *Proc. Roy. Soc.* A, **236**, 119 ; 1956 b, *Times Science Review*, **22**, 17.
MITSUICHI, T., NAGASAKI, H., and UYEDA, R., 1951, *Proc. Imp. Acad., Japan*, **27**, 86.
MIYAKE, S., 1957, *Acta Cryst., Camb.*, **10**, 855 (A).
MOODIE, A. F., 1956 a, *Acta Cryst., Camb.*, **9**, 995 ; 1956 b, *Ibid.*, **9**, 999.
MOLIÈRE, K., and NIEHRS, H., 1954, *Z. Phys.*, **137**, 445 ; 1955, *Ibid.*, **140**, 581.
NAGAKURA, S., 1957, *Acta Cryst., Camb.*, **10**, 601.
PINSKER, Z. G., 1949, *Diffraktsiva Elektronov* (Moscow : Academy of Sciences U.S.S.R.) ; 1953, *Electron Diffraction*, translated by J. A. Spink and E. Feigl (London : Butterworth).
REES, A. L. G., 1953, *J. Roy. Soc. N.S.W.*, **86**, 38.
REES, A. L. G., and SPINK, J. A., 1950 a, *Acta Cryst., Camb.*, **3**, 316 ; 1950 b, *Nature, Lond.*, **165**, 645.
SCHOMAKER, V., and GLAUBER, R., 1952, *Nature, Lond.*, **170**, 290.
STURKEY, L., 1948, *Phys. Rev.*, **73**, 183.
STURKEY, L., and FREVEL, L. K., 1945, *Phys. Rev.*, **68**, 56.
TAKAGI, S., and SUZUKI, T., 1955, *Acta Cryst., Camb.*, **8**, 441.
THOMSON, G. P., and COCHRANE, W., 1939, *Theory and Practice of Electron Diffraction* (London : Macmillan).
WEINSTEIN, B. K., 1949, *C.R. Acad. Sci. U.R.S.S.*, **68**, 301 ; 1952, *Ibid.*, **83**, 227 ; 1953, *J. Exp. Theor. Phys.*, **25**, 157 ; 1955 a, *C.R. Acad. Sci. U.R.S.S.*, **104**, 537 ; 1955 b, *J. Phys. Chem., Moscow*, **29**, 327 ; 1956 a, *Kristallografiya*, **1**, 17 ; 1956 b, *Ibid.*, **1**, 150 ; 1956 c, *Structural Electron Diffraction* (Moscow : Academy of Sciences, U.S.S.R.).
WEINSTEIN, B. K., and PINSKER, Z. G., 1949, *J. Phys. Chem., Moscow*, **23**, 1058 ; 1950 a, *Ibid.*, **24**, 432 ; 1950 b, *C.R. Acad. Sci. U.R.S.S.*, **72**, 53.
WILSON, A. J. C., 1949, *Acta Cryst., Camb.*, **2**, 318 ; 1952, *Ibid.*, **5**, 318.
YAMZIN, I. I., and PINSKER, Z. G., 1949, *C.R. Acad. Sci. U.R.S.S.*, **65**, 645.
ZACHARIASEN, W. H., 1934, *Z. Kristallogr.*, **88**, 150 ; 1954, *Acta Cryst., Camb.*, **7**, 305.

THE SPECIFIC HEATS OF METALS AT LOW TEMPERATURES

By D. H. PARKINSON

Royal Radar Establishment, Great Malvern, Worcestershire

CONTENTS

	PAGE
§ 1. Introduction	227
§ 2. The lattice specific heat	228
2.1. Early development and Debye theory	228
2.2. Lattice spectrum calculations	231
2.3. General considerations	239
2.4. Experimental results	240
§ 3. The specific heat of conduction electrons in metals	242
3.1. Theoretical background	242
3.2. The monovalent metals	245
3.3. The divalent and non-transition metals	248
3.4. The transition metals	251
§ 4. The specific heat of superconductors	255
§ 5. The specific heat of the rare-earth metals	257
§ 6. Experimental techniques	262
§ 7. Conclusion	267
References	268

Abstract. The specific heats of metals at low temperatures have been studied extensively both theoretically and experimentally. There are many processes which contribute to the specific heat, most of them simultaneously which makes the experimental determination of any one difficult. The specific heat is insensitive to the ' finer structure ' of matter and so is of limited value on its own in assessing the refinements of theory. Nevertheless its calculation requires the minimum of theory and it is the ' basic parameter ' which must be given by any general theory describing the energy levels of a crystal.

The contribution due to the excitation of the lattice vibrations is usually represented in terms of θ_D, the Debye characteristic temperature. There have been several attempts to calculate the lattice contribution to the specific heat for metals. In such calculations the chief difficulty is to take account of the effect of the free electron gas on the elastic constants. The free electrons themselves contribute a term which is linear with temperature and which can be observed at temperatures of a few degrees only. Band theory gives a qualitative description of the behaviour of the free electrons in metals but for the complex metals apart from aluminium the quantitative description is vague. The general properties of the transition metals can be explained using the idea of a d-band in which there is a comparatively high density of states and for which the density of states at the Fermi-level can be estimated from the linear ' electronic ' term in the specific heat.

With superconductors the thermodynamic equations governing the transition yield expressions for the free electronic specific heat in the normal state which depend on magnetic parameters only. This gives a convenient way of determining the electronic specific heat but one which is difficult to apply as it requires strain-free specimens of the highest purity. Recently it has been shown that in the superconducting state the specific heat follows an exponential law rather than a cubic with temperature as was originally believed.

The rare earth metals show a great number of spectacular anomalies of magnetic origin in their specific heat curves and include both ferro- and antiferromagnetism. There has been no general theoretical attempt to explain their complex behaviour.

Experimentally specific heats are difficult to determine with precision and although thermal isolation can be achieved comparatively easily there are many sources of error. Below 1°K there has been little systematic work and much remains to be done.

§ 1. Introduction

A discussion of the specific heats of metals must inevitably cover a very wide field, since most temperature dependent phenomena contribute to the specific heat. Characteristic ' anomalies ' occur with ferro- and antiferromagnetics at their Curie and Néel points respectively and with superconductors at their transition temperatures. On a less spectacular scale the free electrons make a small contribution which can be measured at low temperatures and which is related to the electronic band structure of the metal. The crystal lattice itself usually provides the major contribution to the specific heat with which these other effects are mixed and from which they must be extracted for purposes of analysis. In order to limit this discussion ferro- and antiferromagnetism are not dealt with in any detail and alloys are introduced only where they have direct bearing on the pure metals. The lattice specific heat and the influence of the free electrons on it are discussed first. The contribution of the free electrons and its relation to band theory for the simple and transition metals are considered next. Then follow sections devoted to superconductors and the rare-earth metals ; the latter provide a remarkable display of specific heat anomalies. Finally, a section is devoted to some remarks on experimental techniques.

There has been a great deal of activity in the field of low temperature calorimetry in the last decade and most metals have been examined in reasonable states of purity at temperatures extending into the helium range. A number of anomalies, those originally reported in sodium, mercury, beryllium, germanium and hafnium, which have been very puzzling have now been shown not to be characteristic of pure metals. In their place, however, the rare-earth metals provide a greater variety of ' anomalies ' but these are accompanied by other effects and have plausible explanations. The behaviour of pure metals is fairly well understood. Attention has now turned to obtaining more precise data so that the finer points of theory can be settled, and to improving upon the present loose agreement between theory and experiment.

The electrons and the lattice of a metal are often treated as independent systems. This is only a first approximation, for the electron–lattice interaction can be appreciable ; cf. the theories of superconductivity (§ 4). The specific heat has a useful place in elucidating the behaviour of either the lattice or the electrons in the metal. Whatever theory of the solid state is developed it must in the first place lead to a set of energy levels which the ' particles ' of the crystal can occupy. By using suitable statistics it must be possible to find the total energy of the system and by differentiation with respect to temperature find the specific heat. Such a calculation requires the minimum of theoretical structure, which is both the strength and weakness of specific heat studies. The specific heat can provide an effective first test of a theory, but because it is a measure of a mean quantity it does not reveal the fine structure of matter.

In an experiment the specific heat at constant pressure c_p is the quantity measured, and the atomic heat C_p is defined as $\lim_{\Delta T \to 0}(\Delta\theta/\Delta T)_p$, where $\Delta\theta$ is the quantity of heat taken up by the metal per gramme atom and ΔT is the corresponding temperature rise. In theoretical analyses the quantity calculated is usually the specific heat at constant volume for a fixed number of particles,

usually Avogadro's number, i.e. the atomic heat C_v. C_p and C_v are connected by the well-known thermodynamic equation

$$C_p - C_v = \beta^2 VT/K$$

where β is the volume coefficient of expansion, V the atomic volume and K the compressibility. This difference at low temperatures is usually very small and will not be referred to again.

It is well known that classical theory leads to Dulong and Petit's law, $C_v = 3R$, when applied to the calculation of the lattice specific heat. Similarly, with the concept of free electrons acting like a gas within the metal, it leads to a value of the electronic specific heat $C_{v(\text{elect})} = 3R/2$, much larger than that observed experimentally. This discrepancy and the fact that the specific heats of solids decrease towards zero with decreasing temperature cannot be explained in classical terms. Einstein (1907) took the crucial step of applying quantum ideas to lattice theory and Sommerfeld (1928) first applied them to free electrons in metals. All of what goes before and much of the subsequent early work will be omitted except for the essential ideas.

§ 2. The Lattice Specific Heat

2.1. *Early Development and Debye Theory*

Immediately after the publication of Einstein's original paper the idea that there was a wide spectrum of lattice waves ranging from ' audio frequencies ' right up to ' infra-red frequencies ' took shape (see, for example, Madelung 1909 and Sutherland 1910). It was during this period that the current way of regarding lattice specific heats developed. If the number of permitted lattice waves in one gramme atom whose frequencies lie between ν and $\nu + d\nu$ is $q(\nu)\,d\nu$, the total number of vibrations is $\int q(\nu)\,d\nu$. Clearly this number cannot exceed $3N$, where N is Avogadro's number, and there must be some maximum frequency such that

$$\int_0^{\nu_m} q(\nu)\,d\nu = 3N. \qquad \ldots\ldots(1)$$

Since the lattice waves obey Bose–Einstein statistics the mean energy of the crystal is

$$E = \int_0^{\nu_m} q(\nu)\,h\nu\,\{\exp(h\nu/kT) - 1\}^{-1}\,d\nu, \qquad \ldots\ldots(2)$$

and on differentiation with respect to T equation (2) gives the atomic heat. The problem is to find a satisfactory method of calculating $q(\nu)$.

Einstein (1907) assumed that all $3N$ ' oscillators ' had the same frequency, the spectrum of figure 1 (*a*), which leads to the Einstein expression for specific heats given in elementary text-books. This equation does not hold over very extended temperature ranges.

Nernst and Lindemann suggested as an empirical measure that the spectrum consisted of two lines, the one having a frequency half that of the other, figure 1 (*b*). This leads to the expression

$$C_v = \frac{3R}{2}\left[f_E\left(\frac{\theta}{T}\right) + f_E\left(\frac{\theta}{2T}\right)\right] \qquad \ldots\ldots(3)$$

where θ and $\theta/2$ are the characteristic temperatures for the Einstein function f_E. This gives a much better fit to experimental results and is a successful empirical formula.

Debye (1912) assumed that the crystal lattice can be represented by a continuous elastic medium and calculated the normal modes of vibration assuming cyclic boundary conditions. Allowing for the fact that there are both transverse

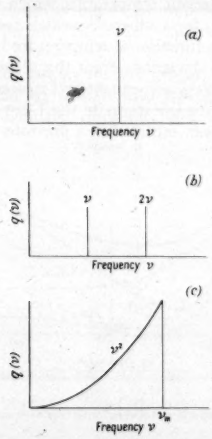


Figure 1. Comparison between (*a*) Einstein, (*b*) Nernst and Lindemann and (*c*) Debye vibration spectra.

and longitudinal modes of vibration and that each transverse mode must be considered as two independent modes at right angles to each other, his full expression for $q(\nu)$ was

$$q(\nu)\,d\nu = \frac{4}{3}V\left(\frac{1}{C_l^3}+\frac{2}{C_t^3}\right)\nu^2\,d\nu = \text{constant}\ \nu^2\,d\nu \qquad \ldots\ldots(4)$$

where C_l and C_t are the velocities of longitudinal and transverse waves respectively, and V is the volume of the crystal.

The form of the spectrum is shown in figure 1(*c*). Using equations (1) and (2) and differentiating the energy lead to Debye's well-known expression for the atomic heat

$$\begin{aligned} C_v &= 9R\left(\frac{kT}{h\nu_m}\right)^3\int_0^{h\nu_m/kT}\frac{e^x x^4}{(e^x-1)^2}\,dx \qquad \ldots\ldots(5)\\ &= 9R(T/\theta_D)^3 f(\theta_D/T) \end{aligned}$$

where $h\nu/kT = x$ and $\theta_D = h\nu_m/k$. $f(\theta/T)$ is the Debye function. At low temperatures equation (5) can be expressed, to within 1%, as

$$C_v = (12R\pi^4/5)(T/\theta_D)^3 \text{ up to about } \theta_D/12.$$

By plotting values of C_v against T/θ_D the results of experiments on many different elements may be superimposed on one curve over a very wide range (Schroedinger 1926); moreover, the T^3 dependence is found at low temperatures. Altogether this theory seemed very satisfactory; it had the simplicity of expressing C_v in terms of a single parameter θ_D and its derivation was easily understood—hence its popularity. It has become customary to discuss the lattice specific heat in terms of a Debye characteristic temperature which may be plotted as a function of temperature. Obviously, for a substance which can be described by the Debye model, θ_D when plotted as a function of temperature is a constant, deviations from this constancy representing deviations from the ν^2 frequency distribution. This method of representation is very convenient and is used in this paper. In the two decades after 1912 various modifications to the Debye theory were suggested and the important points have been outlined in a previous report by Blackman (1941).

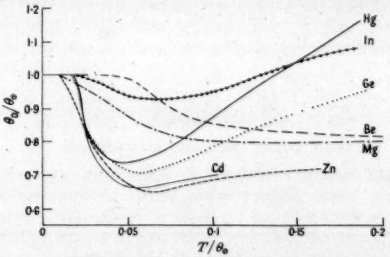


Figure 2. Typical experimental results for θ_D as functions of temperature.

The first difficulty which is encountered with the Debye theory arises from the fact that the velocities of transverse and longitudinal waves used in equation (4) should be calculable from the elastic constants. If the elastic constants measured at room temperature are used for calculating $q(\nu)$, good agreement between calculated and experimental values of θ_D is usually found. However, if values for the elastic constants extrapolated to absolute zero are used, the good agreement disappears (Euken 1913), and the experimental and calculated values of θ_D differ by up to 25% or more (see also Gruneisen and Goens 1924). This sort of disagreement is not surprising and it is remarkable that the Debye theory can fit experimental results so well. The prime criticism is that this model ignores the periodic structure of the crystal which causes it to be a dispersive medium, i.e. the velocity of propagation of the lattice waves is a function of frequency.

As the range of experiment extended towards 1°K the fact that experimentally deduced values of θ_D for most elements are very temperature dependent became apparent. Many examples of such variations have been published and figure 2 shows the results for some metals, the electronic specific heat having been subtracted from the data.

For convenience in getting all the curves in one diagram both θ_D and T have been reduced by dividing by θ_0, the value of θ_D at the lowest temperature. The T^3 region corresponds to the initial horizontal part of the curves.

2.2. *Lattice Spectrum Calculations*

The periodic structure of the crystal must be taken into account in any comprehensive theory. The first step in this direction was taken by Born and von Kármán (1912), who considered a simple cubic Bravais lattice but did not make a detailed study of the spectrum. Their approach to lattice theory was correct but resulted in cumbersome mathematics compared with the Debye theory, and their final expression owing to an unfortunate approximation did not agree very well with experiment. In consequence this work was overshadowed by Debye's.

Born (1923) laid the foundations for the general theory of lattice vibrations, and showed that they could be broken into two groups, the lower frequency group being termed the acoustic vibrations and the higher frequency group

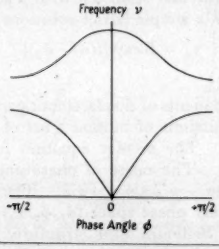


Figure 3. Relation between frequency $\nu = \omega/2\pi$ and phase angle ϕ for diatomic linear chain. The wave number $\sigma = \phi/\pi a$ where a is the atomic spacing.

the optical vibrations. He also showed that for long wavelengths the spectrum obeys a ν^2 law and hence the T^3 variation for C_v follows at low temperatures.

An important step was the publication of a series of investigations by Blackman (1935 a, b, 1937 a, b). By considering a linear chain, a two-dimensional square, and finally a simple cubic lattice, he was able to show that the lattice vibration spectrum is very different from that of Debye, and that the variations in θ_D as a function of temperature such as those of figure 2 can be expected quite naturally.

An analysis of a three-dimensional lattice is inevitably a complex affair. To understand what is likely to happen it is convenient to think of the general diatomic linear chain, having two atoms per unit cell which may differ in mass. Such a chain behaves in exactly the same way as an electrical transmission line. Firstly, it is highly dispersive, i.e. the velocity of propagation of waves depends on the frequency, and secondly, there is a band of frequencies which is not transmitted, below and above which transmitted bands exist. One may plot the transmitted frequencies against wave number (figure 3), and each branch in the diagram corresponds to a different band. Two such branches always occur in the spectrum when there are two atoms per unit cell. This analysis can be extended to take account of the nearest neighbour and further interactions (cf. Brillouin 1946) and also to any number of atoms per unit cell. In general, if there are N atoms per unit cell there are N non-overlapping branches in the spectrum.

In a three-dimensional analysis the force constants between atoms have to be selected. In work published hitherto they reduce to three types. First, simple ' quasi-elastic ' central forces which act along the line joining two neighbouring atoms, and which obey Hooke's law so that they can be characterized by a force constant α. Second, non-central forces arising from the change of angle between a pair of bonds, which were originally introduced by Born. The third type of force is due to the presence of electrons in the metal and has been called a ' volume force '. It can be regarded, to a first approximation, as due to the elasticity of the ' electron gas ', which has a bulk modulus but cannot support a shear stress.

Having chosen a particular model to include the desired forces, the equations of motion of the atoms can be set out. With a simple Bravais lattice there are just three equations for a particular atom but with p atoms in a unit cell there are $3p$ equations per cell. With a simple lattice solutions will be of the form

$$x_n = A\exp\{i(\omega t-\phi_1)\};\ y_n = B\exp\{i(\omega t-\phi_2)\};\ z_n = C\exp\{i(\omega t-\phi_3)\} \quad \ldots\ldots(6)$$

where x_n, y_n, z_n are the components of displacement parallel to the axes of reference. On substitution into the equations of motion a set of simultaneous equations for A, B and C are obtained. The secular equation for the allowed frequencies is then a 3×3 determinant. The range of phase angles ϕ can be limited as in the one-dimensional lattice to $-\pi/2 \leqslant \phi < +\pi/2$. The allowed frequencies can be plotted in a three-dimensional phase space (ϕ_1, ϕ_2, ϕ_3) and the restriction on the ranges of ϕ gives rise to a Brillouin zone structure, just as with the theory of the motion of electrons through a lattice.

In principle, the surfaces of constant frequency can be calculated in the Brillouin zone and $q(\nu)\,d\nu$ is given by the number of modes of vibration lying between the shells of constant frequency in ϕ space. Normally the surfaces of constant frequency cannot be found analytically and suitable numerical methods and approximations have to be employed. Such an analysis gives $q(\nu)$ in terms of the selected force constants.

With his work on a simple cubic lattice Blackman considered quasi-elastic forces using force constants α_1 and α_2 for nearest and next nearest neighbour interactions respectively. He arbitrarily chose $\alpha_2/\alpha_1 = 0{\cdot}05$, which had the advantage of easing the numerical calculations. Using the technique outlined above, he arrived at the spectrum of figure 4 where b, c and d correspond to the three branches of the acoustic waves. From this spectrum a specific heat curve may be calculated and expressed in terms of an equivalent Debye temperature θ_D (figure 5). The curve resembles many which have been found experimentally. Blackman pointed out that the detailed shape depends on the ratio α_1/α_2. Although this model does not correspond to any particular crystal, it provides evidence that considerable variation in the equivalent θ_D as a function of temperature can be expected. At the lowest temperatures, $T < \theta_D/50$, C_v varies as T^3, corresponding to the initial horizontal part of the curve of figure 5. The pronounced minimum at higher temperatures is due to the first maximum in the vibration spectrum where the number of modes is increasing faster than with a ν^2 variation. If the minimum in θ_D is broad and flat it can be mistaken for the true T^3 region in

experiments which go only to hydrogen temperatures (10 to 20°K). This has sometimes been called a ' pseudo-T^3 ' range.

If the results are to be of practical value for a crystal the force constants must be expressed in terms of the macroscopic elastic constants or other parameters accessible to experiment. Simplifying assumptions must be made to reduce the number of force constants so that they can be uniquely determined.

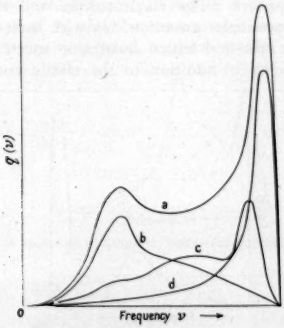


Figure 4. Vibration spectrum of simple cubic lattice (*after* Blackman). Curve a is the total spectrum.

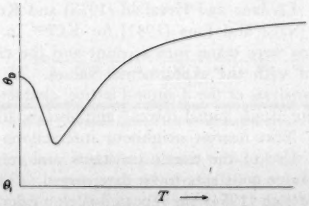


Figure 5. Variation in θ_D for Blackman's simple cubic lattice (scales arbitrary).

Thus in principle the secular determinant for the long wave limit, $\phi \to 0$, for the lattice waves could be compared with that for the propagation of sound waves in the lattice written in terms of the elastic constants. If central forces are used the Cauchy relations must exist between the elastic constants, which immediately results in an apparent difficulty, for with most metals the elastic constants do not obey these relations. Thus with cubic crystals there is one Cauchy relation, namely $c_{12} = c_{44}$, but where data exist for the cubic metals $c_{12} > c_{44}$ often by a factor of two or more. Fuchs (1936) has shown that the most probable cause for the failure of the Cauchy relations for a monovalent metal is due to the effect

of the conduction electrons on the compressibility of the crystal. Those combinations of the elastic constants involved in strains where the volume is not changed should remain unaffected by the electron gas. Thus with cubic crystals c_{44} and $c_{11}-c_{12}$ should be unchanged by its presence ; these are the ' Fuchs conditions ' which should be taken into account in evaluating the force constants of a model representing a metal.

There have been a number of analyses of vibration spectra of various crystals since Blackman's first papers. The alkali halides and the diamond structure elements have received particular attention, for with such crystals useful experimental data such as the infra-red lattice absorption spectra and reststrahlen frequencies are usually known in addition to the elastic constants. In principle,

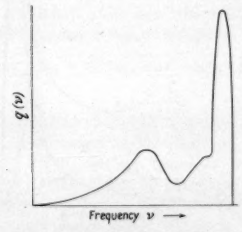


Figure 6. Vibration spectrum of diamond. (*After* Smith.)

it should be possible to establish a much more comprehensive model than is possible with metals. Llydane and Herzfeld (1938) and Kellermann (1941) have made calculations for NaCl and Iona (1941) for KCl. In this work long-range electrostatic interactions were taken into account and the calculated values of θ_D are in good agreement with the experimental values. H. J. Smith (1947) has carried out a detailed analysis of the diamond lattice, characterizing nearest neighbour interactions by the usual ' radial forces ' and also by angular forces as in the original Born model. Next nearest neighbour interactions were represented by simple elastic forces. Use of the elastic constants and reference to the Raman spectrum enabled the force constants to be determined. The resultant spectrum is shown in figure 6. Hsieh (1954) has repeated such a calculation for germanium which gives values for θ_D which vary in the right way but disagree by about 5 to 10% with experiment. A detailed discussion of these analyses is not warranted here ; in general, the results show vibration spectra which have two pronounced peaks in the manner of figures 4 and 7.

The three common crystal structures of metals are body-centred cubic, face-centred cubic, and hexagonal closed-packed. Detailed calculations of vibration spectra have been carried out for the first two only. With cubic crystals the condition for elastic isotropy is that $c_{11}-c_{12}=2c_{44}$ (Wright 1930). This implies that, in general, waves can be regarded as purely transverse or longitudinal in all directions. Tungsten has a body-centred cubic structure and obeys this relation ; it should therefore be comparatively easy to analyse. Fine (1939) published an

analysis considering nearest and next nearest neighbour interactions only, using values of the elastic constants measured at room temperature. To get over the electronic volume effect he employed the Cauchy relation $c_{12} = c_{44}$ and ignored the observed value of the compressibility, assuming it to be influenced by the electrons and not characteristic of the lattice. The vibration spectrum found with this model is shown in figure 7, where two pronounced maxima are prominent features. At very low frequencies a ν^2 distribution was found and C_v should have a T^3 dependence on temperature. Figure 8 shows a comparison between Fine's calculated values of θ_D and the most probable values taken from experiments. The agreement is very good in view of the simplicity of the model.

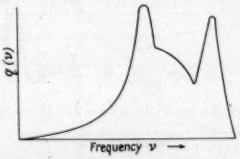


Figure 7. Vibration spectrum of tungsten, body-centred cubic lattice. (*After* Fine.)

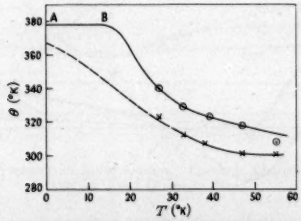


Figure 8. Variation of θ_D for tungsten. Full line, experimental curve, AB due to Waite *et al.*; broken line, Fine's calculated curve.

In his paper Fine compared his results with those of Zwikker (1928) and Lange (1924). These data refer to temperatures above 20°K only but have been extended recently by Waite, Craig and Wallace (1956) to 4°K, and by Daunt and Horowitz (1953) to 2°K (superseding Daunt and Silvidi 1950). The latter results are more difficult to analyse owing to their experimental error and Daunt and Horowitz give a value of $\theta_D = 278$ in the T^3 range, whereas Waite *et al.* find $\theta_D = 378$, which seems to be the more reliable value and is used in figure 8. It is notable that in the experimental results a T^3 term is found right up to 14°K or about $\theta_D/25$.

Leighton (1948) carried out a rather more extensive calculation on the face-centred cubic lattice using Hooke's law forces characterized by force constants α_1 and α_2 between nearest and next nearest neighbours respectively. The vibration spectrum was first determined in terms of the force constants for the ratios $\alpha_2/\alpha_1 = 0$ and $\alpha_2/\alpha_1 = -0{\cdot}1$. Figure 9 shows the complete spectrum for the second

of these ratios. These results were applied to a complete calculation on silver using elastic constants extrapolated to absolute zero. Figure 10 shows the calculated and experimental values of θ_D as a function of temperature. The other curves will be mentioned below. The experimental results shown include the recent work of Corak, Garfunkel, Satterthwaite and Wexler (1955), and Keesom and Pearlman (1955); the older results of Keesom and Kok are ignored below 10°K where they showed a sharp peak in θ_D which has not been subsequently confirmed. The agreement between calculation and experiment is rather good.

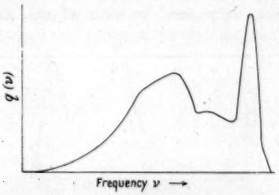


Figure 9. Vibration spectrum of face-centred cubic lattice. (*After* Leighton.)

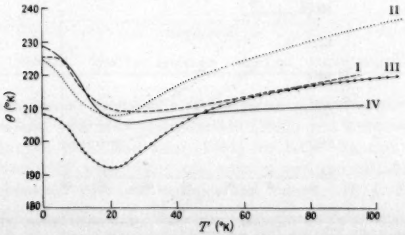


Figure 10. Variation of θ_D for silver. I, experimental; II, calculations of Bhatia and Horton (1955) with elastic constants extrapolated to absolute zero; III, calculations of Bhatia and Horton (1955) with elastic constants for room temperature; IV, calculations of Leighton.

It has been remarked that Fuchs (1936) first investigated the effect of the electronic structure of a metal on the elastic constants, and this point has since received further attention. Thus de Launay (1953, 1954, 1956a) has developed a simple phenomenological model in which the effect of the free electron gas is introduced by attributing to it a bulk modulus K_e, and assuming that it can support no shear stress. Such an effect should influence longitudinal waves only. Where waves are not purely transverse or longitudinal, which is general with non-isotropic metals, displacements can be resolved parallel and perpendicular to the direction of propagation and the electronic effect is assumed to operate in the longitudinal component only. With cubic crystals such a model conforms to the Fuchs conditions and further it can be shown that $c_{12}-c_{44}=K_e$, the electron bulk modulus making up the discrepancy in the Cauchy relationship. De Launay showed that

at the lowest temperatures the calculated value of θ_D for copper, including the effect of the electrons, was in better agreement with experiment than a calculation based on the lattice model of Leighton (figure 11). He suggests that the effect of the electrons in the non-dispersive region of the spectrum may be considerable, whereas in the dispersive part it may be negligible.

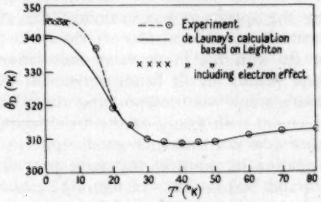


Figure 11. Variation of θ_D for copper.

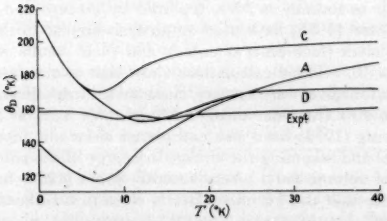


Figure 12. Variation of θ_D for sodium. Curve A, Bhatia's calculation ; Curves C and D due to Bauer.

Bhatia (1955) has carried out a similar investigation for the cubic metals. He assumed that the forces on an ion in the lattice arise from a central interaction $W(r)$ which is significant for nearest neighbours only and also from energies which depend on the atomic volume only and which are due to free electrons, e.g. Fermi, exchange and correlation energies. Examination of the secular equations showed that the results could be expressed in terms of effectively three force constants, the first two depending on the first and second derivatives of $W(r)$ and the third on the presence of the electron gas. This third constant contains two terms : the first corresponds to the bulk modulus of a free electron gas and the second arises from exchange interactions. If the exchange interactions can be neglected, this model is similar to that of de Launay as far as the electrons are concerned. Bhatia applied his results to the calculation of the frequency spectrum of sodium using a method due to Houston (1948). Owing to the mathematical technique involved a number of sharp peaks appear in the frequency spectrum, which in fact go to infinity but have finite areas under them. They group together in a manner not unlike the rounded maxima shown in other analyses. From the spectrum Bhatia calculated the specific heat of sodium and figure 12 shows the

plot of the equivalent θ_D as a function of temperature. In this figure experimental values for θ_D are also given based on recent work by Roberts (1957 a) and Parkinson and Quarrington (1955), for $T < 20°K$, and the older results of Simon and Zeidler for the higher temperature range. The variation of the experimental values of θ_D for sodium is remarkably small. Bauer (1953) used a two-constant model and ignored the effect of electrons within the material; he also used Houston's method for calculating the spectrum but examined the effect of flattening the spectrum in two different ways. His results are shown in figure 12 also and the most likely curve for θ_D with his methods of calculation should lie between curves C and D. These results are in better agreement with experiment above 40°K than those of Bhatia's analysis incorporating the effect of electrons, which gives rather better agreement with $\theta_{D(expt)}$ at the lowest temperatures. The more recent experimental work does not lend such good support to de Launay's suggestion of electron participation in the non-dispersive part of the spectrum as the older results of Pickard and Simon did. In figure 12 experimental θ_D values are not given above 40°K since a small variation in C_v corresponds to a large variation in θ_D at higher temperatures. Note that there is no confirmation of the anomaly in the specific heat in sodium at 7·5°K reported by Pickard and Simon (1948).

Bhatia and Horton (1955) have used an analysis similar to that of Bhatia outlined above with silver (face-centred cubic), and their results have already been displayed in figure 10. These calculations show that at absolute zero the theory of Bhatia and Horton gives results very close to the experimental values of θ_D, while above 30° to 40°K the older theory of Leighton's is in far better agreement. Leibfried and Brenig (1953) have also considered electronic interactions, using a central force model and assuming the change in energy of the conduction electrons to be a function of volume only. Very recently Jones (1957) has considered the special circumstance that the Fermi surface is close to a Brillouin zone boundary (see § 3.2).

Slutsky and Garland (1957) have made a preliminary study of hexagonal close-packed crystals using a central force model and three force constants, α for the nearest neighbours in plane, β for nearest neighbours out of plane and γ for next nearest neighbours out of plane. They have shown how the calculations may be carried out and have pointed out the value of high speed computers to enable higher precision and greater mathematical detail to be used than has been possible hitherto. However, as no elastic data appear to exist for hexagonal metals at very low temperatures, the results have not yet been applied to a real crystal.

The variation of θ_D as a function of temperature near absolute zero has been considered by de Launay (1956b) and also by Barron and Morrison (1957). These authors point out that deviations from a strictly T^3 variation must occur as the temperature rises and contributions to the specific heat from terms in T^5 and higher powers can be expected above about $\theta_D/50$. Where quasi-T^3 dependence is found at higher temperatures neglect of these contributions could lead to errors both in the estimation of θ_D and the electronic specific heat.

The enumeration of the investigations given above is not intended to be exhaustive. A great deal of work has been done by Montroll (1943) and Montroll and Peaslee (1944) primarily on methods of carrying out the complex calculation of the vibration spectrum. Blackman (1956) and also de Launay (1956b) have given

extensive and detailed reviews of lattice theory applied to the calculation of specific heats.

The problem of calculating the vibration spectrum from a measured specific heat curve has also been considered (Montroll 1942, Kroll 1952, Grayson-Smith and Stanley 1950). The chief result from these works is an excellent demonstration of how insensitive the specific heat curve is to details in the vibration spectrum. In order to deduce the spectrum in this way with any reasonable precision, the experimental lattice specific heat would have to be determined with very high precision from the lowest temperatures. There are, of course, other ways of making an experimental determination of the frequency spectrum. Information can be obtained directly from x-ray scattering and also from neutron scattering; the latter technique should be the more powerful.

We are now in a position to make some general observations. First, it is clear that the lattice vibration spectrum of the crystal is complex and usually has at least two major peaks in it, which accounts for the success of the Nernst–Lindemann empirical formula in which the two Einstein frequencies correspond to these peaks. The lattice specific heat can be expected to increase monotonically with temperature. The variations in θ_D plotted as a function of temperature can be quite large. A T^3 region for the specific heat should be found if the temperature is taken below about $\theta_D/50$, where θ_D is the value of θ at high temperature. Both the theoretical results and variations in θ_D that have been observed show that there can usually be a 'T^3' range above the lowest temperatures where θ_D is going through a broad flat minimum, usually in the hydrogen range. An excellent example of this is shown by zinc in figure 2.

Although the theoretical work is developing rapidly, hitherto the lattice models used have assumed simple elastic forces obeying Hooke's law. The idea of complete superposition and no mutual interference of the lattice waves is assumed. In fact, there must be additional anharmonic interactions if thermal equilibrium is to be established. In a complete theory such effects will have to be taken into account.

2.3. *General Considerations*

The lattice specific heat must not be regarded as anomalous when there is a a variation of θ_D with temperature, as was assumed by Simon (1930), who proposed to break the specific heat curve into a Debye and a Schottky function. The latter represents the specific heat due to electrons being excited between two suitably spaced energy levels. Such a Schottky function is also the difference between two Einstein terms, one half the frequency of the other. In consequence the effect of this proposition is equivalent to adding one group of frequencies to a Debye spectrum and removing another group (figure 13); this gives an approximation to an actual spectrum and better agreement with experiment is inevitable under these circumstances.

As a result of his studies, Raman (1941) proposed a different approach to the study of lattice vibrations. He stated that the crystal vibrations are of two sorts, the elastic vibrations on a macroscopic scale which give a continuous spectrum and a second type which are a set of discrete and enumerable monochromatic frequencies in the infra-red region of the spectrum. In finding the

thermal energy of the solid he considers only the second set of frequencies which then give rise to a set of Einstein terms for the specific heat. A number of analyses have been carried out by Raman's pupils, based on these concepts. However, the whole theory is not based on an *ab initio* mathematical analysis as is the lattice theory of Born and Blackman, but rather on a statement of a method of breaking up the spectrum into a number of Einstein terms. With several disposable constants there is bound to be good agreement with experiment over a wide range of temperature.

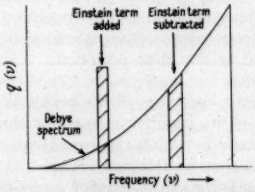


Figure 13. Vibration spectrum corresponding to the use of a Debye and Schottky function for the specific heat.

There have been a number of methods proposed for deducing values of θ_D, such as by examination of the infra-red absorption of the alkali-halides or from X-ray data. Although the values so obtained must be related to the vibration spectrum, the relation is by no means necessarily a simple one. Thus it is possible to deduce values for the Debye temperature from the temperature dependence of the electrical conductivity of a metal. Such a derivation is based on the Bloch theory of conductivity where the longitudinal lattice waves are assumed to interact with electrons to give the resistance. It follows that the θ-values so obtained should be compared with those derived by considering only longitudinal waves as the source of specific heats (Blackman 1951). Such a comparison does not give a good result and serves to demonstrate the complexity of electron lattice interactions.

The use of θ_D to represent specific heat curves has become widespread and a number of ways of comparing elements from the same group have been used. Thus Hill and Parkinson (1952) have plotted θ/θ_x as a function of T/θ_x where θ_x is the value of θ_D at high temperatures. Bijl (1957) has plotted as a function of T/θ_0 the ratio of the experimental specific heat to a Debye specific heat corresponding to the Debye temperature θ_0 found in the low temperature T^3 range. This is a very neat way of comparing results and shows the deviations from the Debye law very clearly. The ratio $C_{expt}/C(\theta_0/T)$ is unity at the lowest temperatures and will tend to unity at the highest temperatures if the specific heat tends to the classical value $3R$.

2.4. *Experimental Results*

A number of experimental results for θ_D have already been displayed in figure 2, as well as curves for silver, copper and sodium in figures 10, 11 and 12. Some recent results due to Roberts (1957) for the alkali metals are shown in

figure 14. At the lowest temperatures the specific heats can be expressed in the form $\gamma T + AT^3$. The cubic term corresponds to the lattice and the linear term corresponds to free electrons (§ 3.2). In calculating θ_D the linear term has been found and subtracted from the results.

Smith (1955) has published a number of results for hexagonal crystals, and although the variations in θ_D are larger than for tungsten, for example, they are not very different from those for germanium (diamond, cubic). From these curves it is evident that the variations in θ_D can be extensive, and although there is the same general trend, there is not necessarily a similar shape in the (θ_D, T) curves, for they simply reflect the coarse structure of the vibration spectrum which depends on the atomic force constants. It is through these that any effect due to crystal structure itself will be produced.

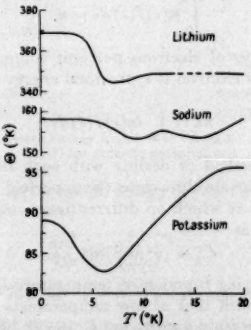


Figure 14. Debye temperatures for the alkali metals. (*After* Roberts 1957.)

The anomalies reported by Cristescu and Simon (1934) in germanium and hafnium have been referred to in the introduction. Pure germanium does not show any such unusual behaviour as has been shown by the measurements of Hill and Parkinson (1952) and Estermann and Weertman (1952). The anomaly in sodium has been referred to above and in § 2.2. Another was reported in beryllium at 11°K, but this also has not been found again by Hill and Smith (1953), who used material of good purity. Similarly, Pickard and Simon (1948) reported anomalous behaviour in mercury at about 7 to 8°K. Smith (1955) was unable to confirm this behaviour but found that θ_D was still varying rapidly even at 3°K. Hafnium was investigated by Friedburg (1954) and no large anomaly found.

In a few instances there has been evidence for believing that the (θ_D, T) curve showed a maximum, as with the results for sodium (figure 14) where there is a small one. This, however, is just outside the estimates of experimental error and should be verified before too much significance is attached to it; cf. silver, for which the original work of Keesom and Kok showed a maximum in θ_D at around 5°K which has not been found subsequently.

§3. The Specific Heat of Conduction Electrons in Metals

3.1. *Theoretical Background*

Many of the basic ideas concerning conduction electrons in metals are parallel to those already outlined in the sections dealing with lattice vibrations. It can be supposed that there are a set of energy levels available to these electrons such that the number lying between energies ϵ and $\epsilon+d\epsilon$ is $n(\epsilon)\,d\epsilon$. Using the Fermi–Dirac distribution function the number of states lying between ϵ and $\epsilon+d\epsilon$ which are occupied is

$$dn = \frac{n(\epsilon)}{\exp\left[(\epsilon-\epsilon')/kT\right]+1}\,d\epsilon = n(\epsilon)f(\epsilon)\,d\epsilon \qquad \ldots\ldots(7)$$

where the Fermi energy ϵ' is defined by

$$\int_0^\infty n(\epsilon)f(\epsilon)\,d\epsilon = n,$$

where n is the total number of electrons per unit volume.

If the mean energy per electron is $\bar{\epsilon}$, the total energy is

$$n\bar{\epsilon} = \int_0^\infty \epsilon n(\epsilon)f(\epsilon)\,d\epsilon. \qquad \ldots\ldots(8)$$

There is a well-known method of dealing with such integrals for the condition $kT \ll \epsilon_0'$, the value of ϵ' at absolute zero (Sommerfeld and Bethe 1934). This leads to an expression for $n\bar{\epsilon}$ which on differentiation with respect to T gives the electronic specific heat C_e as

$$C_e = (\pi^2 k^2/3)\,n(\epsilon')\,T. \qquad \ldots\ldots(9)$$

Since $n(\epsilon')$ is a slowly varying function of temperature for $kT \ll \epsilon_0'$, the result (9) is the basis for the statement that at low temperatures the specific heat of free electrons is proportional to temperature, i.e. $C_e = \gamma T$. Further, γ is proportional to the density of states at the Fermi surface, i.e. where $\epsilon = \epsilon'$. Equation (9) is perfectly general and the theoretical problem is to find the form of $n(\epsilon)$ and in particular $n(\epsilon')$.

The free electron model for metals gives a simple way of dealing with some of their properties. The crystal is regarded as a rectangular potential well in which the electrons are free to move. The effects of other electrons can be taken as smeared out over the crystal to give a simple addition to the depth of the well but otherwise ignored. The Schroedinger equation for such a 'box' is $\nabla^2\psi + (2m/\hbar^2)E\psi = 0$. The boundary conditions can be taken as cyclic, cf. Debye theory. This is a well-known 'problem' and dealt with in text-books. Solutions are of the form $\psi = e^{i\mathbf{k}\cdot\mathbf{r}}$ where $\mathbf{k}$ is a wave number vector. It is then easy to show that the relation between energy and $\mathbf{k}$ is

$$E = (\hbar^2/2m)\,\mathbf{k}^2. \qquad \ldots\ldots(10)$$

This parabolic relation is shown as the curve in figure 15. It is convenient to think in terms of a three-dimensional space for $\mathbf{k}$ which is obviously related to the reciprocal lattice. Surfaces of constant energy in $\mathbf{k}$ space correspond to surfaces of constant frequency in the similar σ-space or phase space when dealing

with lattice vibrations. Equations such as (10) represent a set of spheres (cf. Debye) and it is quite easy for this simple example to sum the energy states lying between shells corresponding to ϵ and $\epsilon+d\epsilon$. The result is

$$n(\epsilon)\,d\epsilon = \frac{1}{2\pi^2}\left(\frac{2m}{\hbar^2}\right)^{3/2}\epsilon^{1/2}\,d\epsilon = C\sqrt{\epsilon}\,d\epsilon. \qquad \ldots\ldots(11)$$

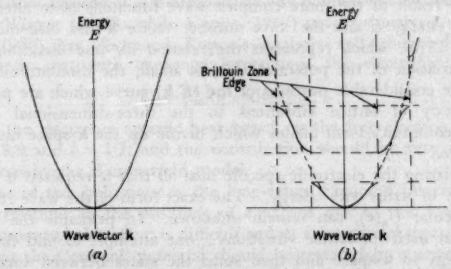


Figure 15. Relation between energy E and wave vector $\mathbf{k}$: (*a*) for free electrons; (*b*) for periodic potential.

This is the form of the spectrum of energy levels for the free electron model. Using equation (9) it follows that

$$C_e = \frac{4\pi^2}{3h^2}k^2\,V\left(\frac{3N}{\pi V}\right)^{1/3} mT = \gamma T \qquad \ldots\ldots(12)$$

$$= 3{\cdot}26\times10^{-5}\,V^{2/3}\,N_a^{1/3}\,T \text{ cal mole}^{-1}\text{ deg}^{-1} \qquad \ldots\ldots(13)$$

where N_a is the number of electrons per atom and V is the atomic volume.

Although this model is over-simplified, the fact that the electrons are moving in a periodic potential field due to the lattice can be taken into account to some extent by giving them an ' effective ' mass m^*, and in place of equation (12)

$$C_e = \frac{m^*}{m}\gamma T. \qquad \ldots\ldots(14)$$

This is equivalent to altering the relation between the energy E and wave number vector $\mathbf{k}$. For the monovalent metals N_a is unity and equation (14) can be compared with the experimental linear term in the specific heat, so giving

$$m^*/m = \gamma(\text{expt})/\gamma(\text{free electron}).$$

It is immediately possible to carry out an order of magnitude calculation to compare the likely values of the electronic specific heats with those of the crystal lattice which shows that the electronic contribution is only likely to be easily observable at below 4°K for simple metals.

An obvious improvement in the theory is to include the effect of the periodic potential $V(\mathbf{r})$ due to the lattice in the wave equation

$$\nabla^2\psi + (2m/\hbar^2)[E - V(\mathbf{r})]\psi = 0. \qquad \ldots\ldots(15)$$

Bloch's well-known theorem states that the solutions of this equation are of the form

$$\psi = U_{\mathbf{k}}(\mathbf{r})\, e^{i\mathbf{k}.\mathbf{r}} \qquad \ldots\ldots(16)$$

where $U_{\mathbf{k}}(\mathbf{r})$ has the periodicity corresponding to $V(\mathbf{r})$. Obviously the special case where $U_{\mathbf{k}}(\mathbf{r})$ is constant corresponds to the Sommerfeld model. Bloch's theorem is fundamental in a theoretical study of the behaviour of electrons in solids. The result of the more complex wave functions is to alter the relation between the energy E and the wave number vector $\mathbf{k}$ and discontinuities occur as in figure 15 (b), which represents the relation for one direction only. Provided the variation in the potential $V(\mathbf{r})$ is small, the discontinuities are small and there are considerable portions of the $(E, \mathbf{k})$ curve which are parabolic. As with the study of lattice vibrations in the three-dimensional σ-space, the points of discontinuity form planes which divide up the $\mathbf{k}$-space into zones—the Brillouin zones.

For calculating the electronic specific heat all that is required is the variation of the density of states with energy. The exact form of the wave functions (15), and in particular $U_{\mathbf{k}}(\mathbf{r})$, can remain unknown. In principle the procedure is similar to that used for lattice vibrations; one attempts to find the surfaces of constant energy in $\mathbf{k}$-space and then sums the states between successive shells $d\epsilon$ apart. As the model taken for the solid becomes more complex, this problem becomes mathematically more difficult and a variety of methods have been devised for dealing with it. It is out of place to discuss them here since there are reviews by Reitz (1955) and Raynor (1952).

The consequence of the discontinuities in energy at the Brillouin zone boundaries is that, on a (density of states, energy) diagram, the permitted states fall into a set of bands, one to each zone, which in metals usually overlap. If the atoms of a metal are imagined to be separated slowly, the discontinuities in the $(E, \mathbf{k})$ curve of figure 15(b) become more marked, and as the wave functions of individual atoms overlap less and less the energy becomes less dependent on $\mathbf{k}$ until eventually the diagram of figure 15(b) becomes simply a set of discrete horizontal lines. In other words, the density of states corresponding to any small energy interval in a particular band increases and eventually the bands degenerate into a set of levels corresponding to those of the free atom. As this goes on, the character of the functions $U_{\mathbf{k}}(\mathbf{r})$ changes until they correspond to wave functions of the free atoms.

With the alkali metals there is one s electron outside closed shells. The closed shells are comparatively undisturbed and the band of energies available to the conduction electrons is made up from the overlapping s wave functions. Consequently, there are $2N$ states available in the s band for the N electrons in each gramme atom and the first zone is therefore only half filled in these metals. Provided the energy surface in $\mathbf{k}$-space corresponding to the Fermi energy does not get too near the zone boundaries it should be a reasonable approximation to take it as a sphere and the free electron model should apply (cf. table 1). This point is discussed further in § 3.2. For the divalent metals there are two electrons per atom to be accommodated in the band system. The first zone should be full, but in fact the second zone overlaps the first to some extent; consequently, the first is nearly full and the second partially filled. The Sommerfeld model cannot

be expected to yield good results for these and the remaining more complex metals.

In the theories of ferromagnetism and of the cohesive energies of metals exchange interactions are important. One would expect therefore that their inclusion in the theory would improve the agreement between the calculated and experimental electronic specific heat. However, when exchange terms are introduced, qualitative agreement is destroyed, as was pointed out by Bardeen (1936), and also by Wigner (1947), while Koppe (1947) gave an analysis of this point. Wohlfarth (1950) has discussed Koppe's analysis and amended his results. Wohlfarth finds that with exchange interactions the electronic specific heat should be

$$C_e = C_{ef}/(a - b\log_{10} T) \qquad \ldots\ldots(17)$$

where C_{ef} is the electronic specific heat based on the free electron model. For sodium $a = 8.9$ and $b = 1{\cdot}7$, and the contribution should be very much smaller than that given by the Sommerfeld model.

The cause of this behaviour is the long-range Coulomb interaction and the effect of correlations in the positions of electrons due to this interaction. An accurate treatment of this effect is difficult and as an empirical measure Wohlfarth suggested that the Coulomb potential should be replaced by a screened potential $V(r_{12}) = e^2 r_{12}^{-1} \exp(-\lambda r_{12})$. With a value of about 10^8 cm^{-1} for λ, Wohlfarth's expression for the free electron specific heat is about 8% below that of the free electron theory and C_e is again proportional to T. Landsberg (1948) found it necessary to use a screened potential in considering the theory of the tail of the soft x-ray emission curve of sodium. Lee-Whiting (1952) has also used such a term in calculations of x-ray absorption line widths. He gives a method of calculating λ, assuming that the origin of the exponential screening lies in " the polarization of the electron gas around each electron " and finds values of λ which agree well with those used by Landsberg and Wohlfarth.

Bohm and Pines (1951, 1952, 1953) have developed a method for dealing with the motions of the electrons within a metal in which the Coulomb interactions and correlated motions are included at the outset. They are able to show that the short-range interactions are equivalent to those of free particles interacting with a screened potential and that the long-range interactions can be neglected. This development removes the difficulties referred to above, accounts for the success of the Sommerfeld model and justifies the use of a screened potential in the x-ray calculations.

3.2. *The Monovalent Metals*

The results of experiments on the monovalent metals are now considered. The alkali metals, particularly sodium, are favourite subjects for calculations because they are ostensibly very simple metals. They have not received as much attention experimentally as could be desired owing to the difficulties of handling them. Results for the helium range of temperatures have recently been published. Table 1 shows the values of γ_{expt}, the coefficient of the linear term in the low temperature specific heat, deduced using the conventional plot of C_p/T against T^2.

In both (1) and (2) of table 1 the experimental range was between 1 and 20°K. The variations in θ_D from (1) are shown in figure 14. In these two papers there

was no evidence of the anomaly in sodium reported by Pickard and Simon (1948). However, with this metal there is a systematic disagreement between these two sets of results of between 5 and 10% over the whole temperature range which is outside experimental error. The reason may lie in a partial crystallographic transformation reported by Barrett (1956), in which a fraction of the metal goes over to a hexagonal close-packed structure from the normal body-centred cubic at about 30°K. The fraction changing depends upon the state of strain and the rate of cooling of the samples which were prepared differently in the experiments of Roberts (1957a) and Parkinson and Quarrington (1955). It appears that sodium is more complex in behaviour than originally thought.

Table 1. Electronic Specific Heats of the Alkali Metals

Metal	γ_{expt} (10^3 joule mol^{-1} deg^{-2})	Ref.	$(m^*/m)_{expt}$	$(m^*/m)_{calc}$ (Pines (3))
Li	1·75	(1)	2·32	0·84
Na	1·37	(1)	1·22	0·90
	2·0	(2)	1·6	
K	1·97	(1)	1·1	0·93
Cs	—	—	—	0·95
Rb	—	—	—	0·96

References: (1) Roberts (1957a); (2) Parkinson and Quarrington (1955); (3) Pines (1955).

The values of γ_{expt} for lithium and sodium have apparent probable errors of about 3%. With potassium, owing to the low value of θ_D and the consequent small range in which a T^3 variation for the lattice specific heat can be observed, the result is much less reliable. Moreover, a certain amount of trouble due to desorption of helium exchange gas was observed below 2°K. Unfortunately, caesium and rubidium are difficult to handle and obtain in good purity and there are no experimental data for them. In view of their low Debye temperatures, work on these metals will have to be carried out below 1°K for reliable values of γ_{expt} to be obtained. Rayne (1954) measured the specific heat of sodium in this very low temperature range using the usual adiabatic demagnetization techniques. He found evidence for a large anomaly around 0·5°K. This should be treated with reserve, at present, for it is just in the range where films of adsorbed helium gas can be very troublesome. In view of these things and the discrepancy between calculated and experimental values of m^*/m shown in table 1, there is urgent need for more work on the alkali metals in the temperature range below 1°K where the lattice specific heat becomes really negligible compared with the electronic.

The interesting feature in table 1 is the discrepancy between the calculated and experimental values of m^*/m, particularly for lithium. If the high experimental values of the electronic specific heat are accepted as correct, a theoretical explanation is required. Buckingham and Schafroth (1954) have suggested that this is due to another aspect of the lattice–electron interaction. In the usual electronic band structure calculations the lattice is regarded as a set of fixed ions, whereas, in fact, the lattice vibrations should be taken into account. An electron must have a 'self-energy' in the vibrational field for it produces deformations of the lattice in its motions which in turn react upon the electron. In the metallic

state this interaction must be considered for all the electrons together. Fröhlich (1950) developed these ideas and discussed this type of interaction and the possibility of its leading to superconductivity. He introduced an interaction parameter F such that if $F > F_0$, a critical value, the metal should be a superconductor and if $F < F_0$ the metal should be a normal conductor. Buckingham and Schafroth have discussed this second situation and have shown that there should be an increase in the electronic specific heat corresponding to an increase in the density of states at the Fermi level (see also Buckingham 1951). It is interesting to note that in Fröhlich's original paper lithium is the one monovalent metal which disobeys his criterion, showing a value for F larger than F_0, and experimentally it shows a remarkably high specific heat. In view of the soft x-ray data of Skinner (1939, 1940) an increase in the density of states at the Fermi level is to be expected.

The remaining monovalent metals—copper, silver and gold—have been studied extensively for they are among the easiest to work with experimentally as opposed to the alkali metals. With only one free electron per atom, the first Brillouin zone should be half filled and the Sommerfeld model is a reasonable approximation. In table 2 the values of γ_{expt} for these metals are given and the references shown in the caption. With copper, Rayne has recently published an extensive series of measurements. The results for his purest specimen agree very well indeed with those of Corak *et al.*, and in view of the consistency of these independent experiments it is reasonable to assume a value of $0{\cdot}687 \times 10^{-3}$ joule $\text{mol}^{-1}\ \text{deg}^{-2}$ for γ_{expt} giving a value for m^*/m of 1·37. Similarly, with silver these same authors agree well and it is evident that m^*/m should be about 0·95. With gold, on the basis of one set of measurements, m^*/m is 1·16.

From the theoretical point of view the essential difference between the noble metals and the alkalis is that in the former the newly closed shell is a d shell, whereas in the latter it is a p shell. The newly closed d shell is more spread out than the p shell and therefore has to be taken into account in any comprehensive calculation of the relation between the energy and the wave vector. Such calculations were carried out for copper by Jones (1937) and Jones and Mott (1937). In the latter a value for γ of $0{\cdot}69 \times 10^{-3}$ joule $\text{mol}^{-1}\ \text{deg}^{-2}$ was calculated in reasonably good agreement with the experimental values.

More recently, Klemens (1954) in discussing the electrical and thermal conductivities of the monovalent metals has suggested that the Fermi surfaces must be very close to the Brillouin zone boundary, and with the noble metals should actually touch it. The Fermi surface should be considerably distorted also. In this way he could explain why the ratios of the high and low temperature thermal resistivities do not agree with those expected from the Bloch free electron theory. Likewise Jones (1955) has remarked that the thermo-electric data on copper suggests that the Fermi surface must be very near the zone boundary. In a later paper Jones (1957) has shown that for this circumstance it is possible for the electrons to interact with the transverse lattice waves and produce an extra linear term in the specific heat at low temperatures, cf. Buckingham and Schafroth. Under shear stresses the zone boundaries can be imagined to move with respect to the Fermi surface. This is equivalent to an electronic influence on the shear constants of the metal. He used these ideas to explain peculiarities in the specific heats of copper containing small quantities of zinc (Smith, to be published).

Such a mechanism might be used to explain the apparent increase in γ_{expt} for copper with impurity content shown in table 2. Comparatively small quantities of impurities, if in solid solution, might be expected to make small changes in the lattice parameters, and possibly change the filling of the first zone.

Table 2. The Electronic Specific Heats of Copper, Silver and Gold

Metal	γ (10^3 joule mol^{-1} deg^{-2}) Sommerfeld	Expt.	Purity (%)	Refs
Copper	0·502	0·686 (i)	99·999	
		0·670 (ii)	99·9	(1)
		0·670 (ii)	99·9 (annealed)	
		0·799 (iii)	99·0	
		0·74	99·9	(2)
		0·75	99·6	(3)
		0·688	99·999	(4)
		0·72		(5)
Silver	0·644	0·68		(5)
		0·67		(6)
		0·66	99·999	(7)
		0·65		(8)
		0·61	99·8	(4)
		0·61	99·99	(9)
Gold	0·638	0·74		(4)

References: (1) Rayne (1956 a); (2) Kok and Keesom (1936); (3) Estermann, Friedberg and Goldman (1952); (4) Corak, Garfunkel, Satterthwaite and Wexler (1955); (5) Rayne (1954); (6) Keesom and Kok (1934); (7) Keesom and Pearlman (1952, 1955); (8) Yates and Hoare (1957); (9) Rayne (1956 b).

The question of the precise shape and position of the Fermi surface is one of extreme interest. From the experimental point of view, specific heat data are not very helpful for the electronic term simply measures a mean effective mass or mean density of states at the Fermi surface. If the effective mass representation of the relation between ϵ and $\mathbf{k}$ can be retained, the simple equation (10) will be replaced by one in which the mass is expressed as a tensor. This problem could be resolved by studies such as cyclotron resonance, the anomalous skin effect, and the de Haas–Van Alphen effect; the last has not yet been observed in monovalent metals. Chambers (1956) has given a very useful survey of the experimental approach to this type of problem (see also Pippard 1955)†.

3.3. *The Divalent and other Non-Transition Metals*

As the divalent metals of group 2 of the periodic table have two valence electrons per atom, the first Brillouin zone should be filled. In fact, the first and second zones overlap to some extent, which accounts for their conductivity. In consequence, the Fermi surface must intersect the boundaries of the first zone

† *Note added in proof.* In a recent paper Pippard (1957) has shown how studies of the anomalous skin effect can be used to deduce the shape of the Fermi surface. He has applied this technique to copper and gives a detailed analysis for that metal including a calculated value for the electronic specific heat. The result agrees with the experimental values but not too much significance should be attached to this.

so that the energy surfaces must have a complex shape. A mean effective mass does not mean much under these circumstances and it is better to talk in terms of the total density of states at the Fermi surface. In table 3 values of γ_{expt} are shown for these metals. Zinc, cadmium and mercury are of course superconductors, the general behaviour of the specific heat of which is discussed separately in § 4. All the values of γ_{expt} shown here are based on specific heat data.

Table 3. Electronic Specific Heats of the Divalent Metals

Metal	γ_{expt} (10^3 joule mol^{-1} deg^{-2})	n_{expt}/n_{free}	Ref
Be	0·22	0·47	(1)
Mg	1·34	1·33	(2)
Ca	2·73	1·82	(3)
Zn	0·654	0·86	(4)
	0·66	0·87	(5)
St	3·64	2·01	(3)
Cd	0·71	0·73	(6)
Ba	2·7	1·44	(3)
Hg	—		(7)

References: (1) Hill and Smith (1953); (2) Estermann, Friedberg and Goldman (1952); (3) Roberts (1957 b); (4) Keesom and Van der Ende (1932); (5) Daunt and Silvidi (1950); (6) Samoilov (1952); (7) Smith (1955).

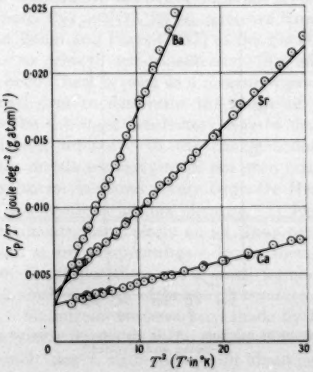


Figure 16. Experimental results for the alkaline earth metals. Plot of C_p/T against T^2.

There are a number of points to be noted concerning the experiments on which the data in table 3 are based. It has already been remarked that beryllium does not show an anomaly at about 11°K (§ 2.4). For the alkaline earth metals the values of γ_{expt} have likely errors of about 3%. Figure 16 shows the usual representation of C_p/T against T^2 for each of them. In measuring the specific heat of mercury Smith (1955) found that there was no anomalous behaviour at 7 to 8°K, as reported by Pickard and Simon (1948). He found, however, that

θ_D appeared to be varying rapidly even at 3°K, and in consequence no reliable values for γ_{expt} could be deduced.

There have been a number of band structure calculations for these metals. Herring and Hill (1940) calculated the band structure of beryllium and found that the 2s and 2p bands overlap. The resulting shape of the density of states diagram is shown in figure 17. According to the calculations, the Fermi level falls just beyond the minimum, corresponding to a ratio of n_{calc}/n_{free} of 0·43, which agrees quite well with the experimental value of 0·47.

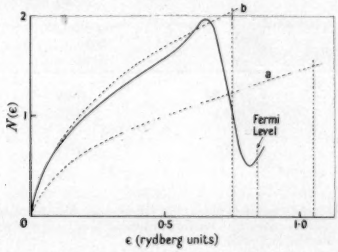


Figure 17. Band structure of beryllium. Curve a for free electrons. (*After* Herring and Hill.)

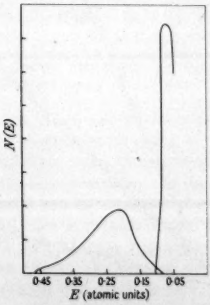


Figure 18. Band structure of calcium. $N(E)$, density of states on arbitrary scale. (*After* Manning and Krutter.)

Manning and Krutter (1937) have made a calculation of the band structure of calcium and find a value of n_{calc}/n_{free} of 1·00 which is very much lower than the value found in experiments. According to the calculations the density of states in the second zone rises sharply and a small error in the calculated overlap could make a considerable difference to the density of states at the Fermi level (figure 18). Strontium and barium also have the face-centred cubic crystal structure and similar electronic configurations in the free atom, therefore they can be expected to have band shapes similar to those in calcium. The degree of overlap in the three metals must vary and presumably be the greatest in strontium.

Of the remaining non-transition metals, all except bismuth are superconductors. They are aluminium, gallium, indium, thallium, tin and lead. Their specific heats have all been measured in the helium range and values of γ deduced, except for gallium. The values of γ_{expt} and references have been tabulated by Keesom and Pearlman (1956). Dolecek (1955) has given a more recent value of γ_{expt} for lead, $3{\cdot}34 \times 10^{-3}$ joule mol^{-1} deg^{-2}; Howling, Mendoza and Zimmerman (1955) find a value of $1{\cdot}37 \times 10^{-3}$ joule mol^{-1} deg^{-2} for aluminium and in table 4 there are extra data for tin. Section 4 deals with the specific heats of superconductors as a special topic.

The electronic band structure of aluminium has received considerable attention. According to calculations by Matyas (1948) the first zone is nearly full and there is approximately one electron per atom in the second. The details of his calculated structure do not agree with results obtained from consideration of the elastic constants (Leigh 1951). Heine (1957 a) has carried out an interesting analysis based on the ideas of Pippard (1955), in which the results obtained from the De Haas–Van Alphen effect, the anomalous skin effect and the measured electronic specific heats are used, together with a fairly simple theoretical calculation of the band structure, to deduce the shape of the Fermi surface. He finds that except for a small region at the corners, containing $3{\cdot}6 \times 10^{-3}$ holes per atom, the first zone is full. There is an overflow of electrons across all zone faces into the second zone and 5×10^{-5} electrons per atom in the third zone near the edges. Heine (1957 b, c) supports this analysis by an extensive band calculation in which he marries the ideas of Bohm and Pines (1953) to the Fock–Hartree equation and in which he attempts to achieve self-consistency. An analysis of this sort in which the electronic specific heat is used as a measured parameter in conjunction with other experimental data to determine the shape of the Fermi surface is valuable and seems to be the most satisfactory way of dealing with these more complex metals. Without reference to experimental data normal theoretical calculations for complex metals are likely to be not even qualitatively correct.

Bismuth is of interest, for it shows a very large De Haas–Van Alphen effect, which was of course first noticed in this element. It has been the subject of many subsequent experiments, particularly on all those parameters which should vary with magnetic field at low temperatures. In the metal two bands just overlap; at the bottom of the upper there are a small number of electrons and in the top of the other a small number of holes. Therefore the electronic specific heat should be small. Measurements have been made by Keesom and Pearlman (1954), who found a small linear term in the specific heat corresponding to $\gamma = 0{\cdot}078 \times 10^{-3}$ joule mol^{-1} deg^{-2}, with an error of about $\pm 0{\cdot}03$. It is considerably larger than would be expected and may very well be influenced by the presence of impurities (sample 99·99% pure only).

3.4. *The Transition Metals*

The transition metals are those in which the unionized free atom contains an incomplete d shell of lower principal quantum number than a completed s shell. In the metal the s electronic wave functions overlap considerably whereas those of the d shell are much less disturbed. Moreover, the d band accommodates ten electrons per atom and the net result is that the density of states in the d band

is considerably higher than the density of states in the s band. The d and s bands overlap but the total density of states at the Fermi level will be governed primarily by that of the d band. It can be expected, therefore, that the transition metals will have coefficients in their electronic specific heat terms which are considerably larger than those of the monovalent metals, for example, which have only an s band partially occupied. This was first pointed out by Mott (1935), who introduced the idea of 'holes' in the d band for metals such as nickel where there is room for about 0·6 electrons per atom in the band.

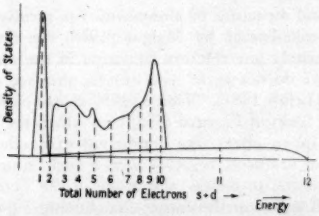


Figure 19. The d band structure of copper. (*After* Slater and Krutter.)

In fact, the transition metals do exhibit a linear component in their specific heat and in many instances high values of γ_{expt}. Krutter (1935) and Slater (1936) carried out calculations of the 3d band for copper; their results are shown in figure 19. As the s band is very shallow, the position of the Fermi surface is governed by the number of electrons which go into the d band. With copper this band is full and makes no contribution to the specific heat but it can be expected that in other elements containing a 3d band, the coarser structure of the band will be similar to that of copper. If then the electronic specific heat for the transition elements in which the 3d shell is being filled is plotted against the total number of s and d electrons, the resulting diagram must show a variation related to the d band structure. This can be done for the 4d and 5d bands also and at least is a very convenient way of showing the data for these metals. Daunt (1955) and Wolcott (1955) have both used such a method, which is also used here (figure 20), which includes new data given by Wolcott. Keesom and Pearlman (1956) have given detailed tabulations of the specific heats of the remaining metals.

It is remarkable that for each of the 3d, 4d and 5d series the trend in the variation in γ is similar, excepting α-manganese for which the high value has been confirmed by Wolcott (1955). For figure 20 to reflect the d band structure with any precision, each element would have to have the same crystal structure with the d wave functions overlapping to the same extent. This clearly is not so, for all the three common metallic structures are represented in each group. Slater and Koster (1954) have calculated the d band structure for a body-centred cubic crystal and their results are shown in figure 21 (*a*). Koster (1955) has carried out similar calculations for face-centred nickel, extending the work of Fletcher and Wohlfarth (1951) and Fletcher (1952), and his results are shown in figure 21 (*b*). The two d bands differ in detail but the principal feature, the

large dip in the middle, is common to both. This coarse structure seems to be shown in figure 20 at $s+d=6$ and is common to all three series. Owing to the changes in crystal structure it is difficult to read more into the results in figure 21 without the aid of other data. Kreissman (1953) has discussed the d band structure using the results of his studies of the high temperature magnetic susceptibilities. The electronic paramagnetic susceptibility should vary as $\chi = \chi_0(1+c\alpha T^2)$

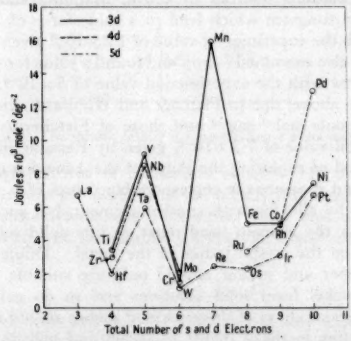


Figure 20. The electronic specific heat constants of the transition elements.

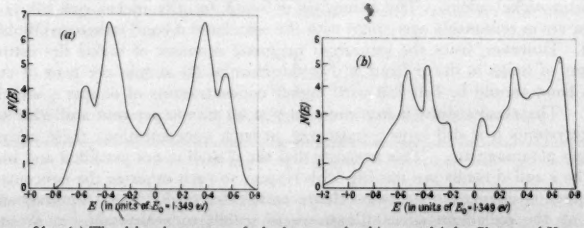


Figure 21. (*a*) The d band structure of a body-centred cubic crystal (*after* Slater and Koster); (*b*) The d band structure of nickel (*after* Koster).

where c is a constant and $\alpha = [\partial^2 \log n(\epsilon)/\partial\epsilon^2]_{\epsilon=\epsilon'}$. The coefficient α alternates in sign systematically through each series. A negative value of α is obtained when the Fermi level is near to a minimum in the density of states curve and a positive value corresponds to its being near a maximum. The alternations in sign correspond to the maximum and minimum in figure 20, from $s+d=4$ to $s+d=7$ inclusive, with a maximum at 7. Kreissman has also discussed the possibility of the alternation in the sign of α and the values of γ being due to antiferromagnetism, but there is not very much evidence to support such an argument. Shull and Wilkinson (1953) have carried out extensive neutron diffraction studies on the transition elements and only chromium and manganese have given evidence of

antiferromagnetism. The Néel point of chromium is at about 475°K and of α-manganese near 100°K. Kreissman and McGuire (1955) have shown that there is a weak maximum in the susceptibility of α-manganese at 125°K and Shomate (1945) has found a corresponding symmetrical λ-like anomaly in the specific heat at 95°K, so for this element the evidence of antiferromagnetic behaviour is very good. The very high value of γ for α-manganese is unexplained.

Besides the calculations referred to above, Manning and Chodorow (1939) made calculations on tungsten which lead to a value for γ of 2×10^{-3} joule mol^{-1} deg^{-2} compared with the experimental value of $1{\cdot}2 \times 10^{-3}$ given by Wolcott (1955). Manning (1943) has also examined α-iron and found a value for γ of $1{\cdot}9 \times 10^{-3}$ joule mol^{-1} deg^{-2} compared with the experimental value of 5×10^{-3}. The calculations on nickel referred to above, due to Fletcher and Wohlfarth, yield a value of γ for nickel of $4{\cdot}7 \times 10^{-3}$ joule mol^{-1} deg^{-2} and those of Fletcher $7{\cdot}1 \times 10^{-3}$, compared with the experimental value of $7{\cdot}3 \times 10^{-3}$, given by Keesom and Clark (1935).

A possible method of exploring the shape of the d band is to alter the position of the Fermi level and measure the corresponding values of γ. The only feasible way of doing this is by alloying with another element, but such a method cannot be applied freely, for the element used must go into solid solution and have an inappreciable effect on the crystal lattice of the metal. Dilute alloys of elements of near atomic number and similar atomic radii are suitable for such purposes. Thus copper and nickel form solid solutions and so do palladium and silver. Values of γ for a series of alloys of these pairs of metals should reflect the variation in the density of states in the d bands of nickel and palladium. Work of this sort was first carried out by Keesom and Kurrelmeyer (1940) with copper–nickel and iron–nickel alloys. The values of γ found for the nickel rich alloys with copper are in reasonable agreement with the calculated d band structure (Wohlfarth 1949). However, from the saturation magnetic moment of nickel the estimated number of holes in the d band is 0·6, therefore at 60 atomic per cent of copper the d band should be full and with higher concentrations of copper γ should be small. The results show a maximum in γ at 60 atomic per cent and with higher concentrations it is still large ; moreover, at such concentrations these alloys are strongly paramagnetic. This suggests that the d shell is not yet filled and implies that the s and d bands are moving with respect to each other as the concentration of copper is changed. Recently Hoare and Yates (1957) have reported similar work on the palladium–silver alloys, where γ falls to a low value at 60 atomic per cent of silver and remains low for higher concentrations.

Iron, nickel and cobalt are, of course, ferromagnetic. In such circumstances, with the energy band representation the d band is regarded as split into two halves, corresponding to oppositely oriented spins, one half moved upwards and the other down on the energy scale. The lower half band falls below the Fermi surface at which the upper only contributes to the density of states. This should be taken into account in interpreting the values of γ in figure 20 in terms of densities of states.

Note added in proof. In a recent discussion of the band structure of the transition metals Mott and Stevens (1957) have considered the body centred metals in particular. In Iron they consider that the magnetic electrons, two in number, do not form a band in the usual way.

§4. The Specific Heat of Superconductors

It is well known that at the superconducting transition there is a very sharp anomaly in the specific heat (figure 22) which in itself is highly interesting. Detailed examination shows that the electronic specific heat of the normal state is intimately bound up with the transition, and moreover can be obtained in principle from purely magnetic measurements. The discussion here will be

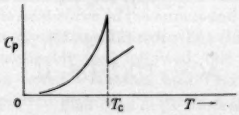


Figure 22. The specific heat of a superconductor at its transition.

limited to just those features of immediate concern. If the specimen can be obtained sufficiently pure physically and chemically, the transition in a magnetic field should be thermodynamically reversible. The equations governing this transition have been studied thoroughly. Using the Clausius–Clapeyron equation

$$\Delta S = S_s - S_n = \frac{V}{8\pi}\frac{dH_c^2}{dT} \qquad \ldots\ldots(18)$$

where S_n and S_s are the entropies of the normal and superconducting states respectively and ΔS is their difference. V is the atomic volume and H_c the threshold magnetic field which represents the equilibrium boundary between the mormal and superconducting phases. It follows directly from (18) that the difference in the specific heat is

$$\Delta C = C_s - C_n = \frac{VT}{4\pi}\left[H_c\frac{d^2H_c}{dT^2} + \left(\frac{dH_c}{dT}\right)^2\right]. \qquad \ldots\ldots(19)$$

At the transition temperature T_c the difference in specific heats becomes

$$[\Delta C]_{T_c} = \frac{VT_c}{4\pi}\left(\frac{dH_c}{dT}\right)^2_{T_c}. \qquad \ldots\ldots(20)$$

These equations have been tested extensively in a number of experiments. Daunt, Horseman and Mendelssohn (1939), for example, showed that a plot of ΔS as a function of T at the lowest temperatures yields a straight line which in agreement with the third law of thermodynamics, goes through the origin; i.e. at 0°K, $\Delta S = 0$. Now it has been demonstrated that in the normal state $C_n = AT^3 + \gamma T$, so that if it can be assumed that any terms in C_s vary more rapidly than linearly, ΔS must tend to S_n as T tends to zero. The temperature dependence of C_s is, in fact, a near approximation to a cubic law (see below). It follows at once that

$$C_n = \gamma T = -\frac{VT}{8\pi}\left(\frac{d^2H_c^2}{dT^2}\right)_{T\to 0}. \qquad \ldots\ldots(21)$$

This equation gives one possible method of obtaining γ from magnetic measurements. An important point is that it is based on the idea of a reversible transition. This is a difficult condition to fulfil experimentally for it demands

the highest possible physical and chemical purity. A second experimental difficulty is that the second derivative of the critical field curve at absolute zero requires accurate experiments taken to the lowest temperatures.

If the specific heat in the superconducting state is assumed to be cubic, it can be shown that the threshold magnetic field should then be parabolic, i.e. $H_c = H_0[1-(T/T_c)^2]$ where H_0 is the critical field at absolute zero. This equation leads to

$$\gamma = VH_0^2/2\pi T_c^2. \qquad \ldots\ldots(22)$$

Similarly, with equation (19) and substituting (22),

$$\Delta C = \gamma T[3(T/T_c)^2 - 1]$$

and at the transition temperature T_c in zero field

$$[\Delta C]_{T_c} = 2\gamma T_c = \frac{VT_c}{4\pi}\left(\frac{dH_c}{dT}\right)^2_{T_c}. \qquad \ldots\ldots(23)$$

There are then the following additional ways of determining γ: (*a*) from the slope of the threshold field curve at $T = T_c$, equation (23); (*b*) from the discontinuity in the specific heat $\gamma = [\Delta C]_{T_c}/2T_c$; (*c*) from equation (22).

Eisenstein (1954) and Daunt (1955) have published reviews of the comparison between γ values determined calorimetrically and from the above equations. In general, reasonable agreement between the magnetic and calorimetric measurements is achieved with the 'soft' superconductors which can be obtained in very pure strain free states. To do this the curve for H_c must be accurately determined and equation (21) used for γ. The remaining methods depend on the assumption of a cubic variation for C_s leading to the parabolic curve for H_c. Table 4 shows a comparison of values for γ for two soft superconductors (tin and indium), and two hard superconductors (vanadium and tantalum).

Table 4. Comparison of γ Values for Soft and Hard Superconductors

$\gamma_{expt} \times 10^3$ (joule mol^{-1} deg^{-2})

Metal	Calorimetric mean values	Magnetic	References for magnetic measurements
Sn	1·82	1·65	Daunt, Horseman and Mendelssohn (1939)
		1·84	Lock, Pippard and Shoenberg (1951)
		1·85	Serin, Reynolds and Lohman (1952)
		1·86	Maxwell (1952)
In	1·81	1·65	Daunt, Horseman and Mendelssohn (1939)
		1·8	Stout and Guttman (1952)
		1·67	Maxwell and Lutes (1954)
V	9·20	6·8	Wexler and Corak (1952)
Ta	5·56	7·95	Eisenstein (1954)
		8·1	Daunt and Mendelssohn (1937)

The results for tin are particularly interesting; Lock *et al.*, Serin *et al.*, and Maxwell were all looking for the variation in T_c with isotopic mass and their experiments were therefore very precise. The agreement amongst the values of γ

from these three sets of measurements is to about 1% and the average value is just more than this above the calorimetric values.

Where the metal is difficult to purify and obtain free from physical defects, as are tantalum and vanadium, the condition of reversibility is not normally met and agreement between the two methods of determining γ is poor. This can be seen from table 4 and is well demonstrated in the reviews of Daunt and Eisenstein already referred to.

The shape of the specific heat curve in the superconducting state could influence values of γ deduced from equations (22) and (23). In earlier measurements a cubic law seemed to be reasonably well followed, but recently deviations from a strictly T^3 law have been noted. Keesom and Pearlman (1956) in their discussion show that C_p/T plotted against T^2 shows a distinct change in slope at about 2·2°c. Similar effects have been reported by Corak, Goodman, Satterthwaite and Wexler (1954), for vanadium and tin and by Goodman (1955) for aluminium. These authors have assumed that the lattice specific heat is unaffected by the transition and is still cubic at the lowest temperatures. In this way the electronic specific heat in the superconducting state could be found. They show good evidence for believing that it follows an expression

$$C_s = \gamma T_c a \exp(-bT_c/T). \qquad \ldots\ldots(24)$$

The difference between the cubic and exponential terms only shows up markedly at the lowest temperatures. Equation (24) was put forward as an empirical relationship, but it is of the form which could be expected if there was a small gap ΔE in the energy level spectrum of superconductors for then at the lowest temperatures a term of the form $\exp(-\Delta E/KT)$ could be expected. The values of a and b are 9·17 and 1·5 respectively for vanadium and tin and 6·9 and 1·28 for aluminium, so the law of corresponding states is not necessarily followed, but $\Delta E \sim kT_c$. Further support for the exponential law has also been produced by Snider and Nicol (1957) in work on thallium. The idea of an energy gap of this order of magnitude has been discussed for many years and recently in particular by Frŏhlich (1954) and Bardeen (1955). There have been several attempts to demonstrate its existence using microwave-surface absorption techniques, such as those of Blewins, Gordy and Fairbanks (1955) and Biondi, Garfunkel and McCoubrey (1956). Tinkham (1956) has extended such work into the far infra-red. He measured the ratio of transmitted power through very thin films of lead in the superconducting state, to power transmitted through normal lead, and found that for a wide band of far infra-red frequencies (500 to 1000μ) the film had a higher transmission than in the normal state. This result can be interpreted in terms of an energy gap $\Delta E \sim 3kT_c$.†

§ 5. The Specific Heat of the Rare-earth Metals

The electronic structure of the atoms of the rare-earth elements can be written as Xe core ; $4f^n$; $5s^2$, $5p^6$; $6s^2$, $5d^1$ with the proviso that a limited amount of interchange between the 4f and 5d shells may take place. As the atomic number

† *Notes added in proof.* (1) Glover and Tinkham (1957) have recently discussed the far infra-red absorption of superconducting films more fully. (2) A very recent and extensive theory of superconductivity has been published by Bardeen, Cooper and Schrieffer (1957).

increases from 57 for lanthanum up to 71 for lutecium, the value of n increases from 0 to 14. The 4f electrons are screened to a large extent by the closed 5s and 5p shells, consequently in these metals this shell should be considerably less disturbed than the d shells in the normal transition metals. If a band scheme is used to represent the energy levels available to the 4f electrons, the density of states in it must be very high and, indeed, such a band must show a very pronounced structure related to the discrete energy levels in the free atom. The salts of the rare-earth metals have been studied extensively using paramagnetic resonance techniques, measurements of susceptibility and related properties at very low temperatures. Magnetic interactions between neighbouring ions and Stark splitting of the energy levels for the 4f electrons due to electric fields in the crystal are important. One can argue by analogy that in the metallic state also, both magnetic and electric interactions are possible and, moreover, are of the same order of magnitude. In consequence, a complicated behaviour depending on the properties of the 4f electrons can be expected with these metals. Nearly all of them show complex magnetic behaviour accompanied by corresponding effects in their specific heat curves. Studies of the specific heats and susceptibilities constitute a powerful combination for mapping the properties of the 4f electrons.

The ions at the lattice points in the metals can be regarded as trivalent with the 6s and 5d electrons going into bands and taking part in the conduction processes. If the screening of the 4f shell is effective the lowest energy states of the trivalent ions may be found using Russell–Saunders coupling and Hund's rules. It then follows that if the conduction electrons are ignored, the susceptibility should follow a Curie–Weiss law with the Curie constant $C = Ng^2J(J+1)\mu_B^2/3k$. Most of these metals follow a Curie–Weiss law at higher temperatures with effective magneton numbers which are remarkably close to those given by this simple formula. The basic idea of using Russell–Saunders coupling for the ions in these metals cannot be far wrong, and this is the only instance where this could occur with metals. Accurate data on a number of the metals are not yet available ; nevertheless, there is enough evidence from the original investigations of Klemm and Bommer (1937), the more recent work of the team led by Spedding *et al.* (1957), and other sources, that as the number of 4f electrons approaches seven, the magnetic interactions increase rapidly. Thus gadolinium ($n = 7$) is ferromagnetic at room temperature. Then as n increases to fourteen, these interactions fall off regularly but are sufficient to make most of the metals in the second half of the series ferromagnetic at the lowest temperatures.

Lanthanum has no 4f electrons, is a superconductor, and behaves in a normal way. The transition temperature is not well defined owing to the difficulty of preparing strain free samples of a single crystalline form ; they are usually mixed face-centred cubic and hexagonal close-packed. It has a comparatively high linear term in the specific heat below 5°K which corresponds to electrons in the 5d band : $\gamma = 8{\cdot}4 \times 10^{-3}$ joule mol^{-1} deg^{-2} (Berman, Zemansky and Boorse 1955, Parkinson, Simon and Spedding 1951). The latter also report measurements on cerium, praseodymium and neodymium.

Cerium has one 4f electron and shows a complex behaviour. With face-centred cubic specimens there is an extensive region between 40° and 180°K in

which there are very large thermal hysteresis effects. On cooling, the specific heat follows the curve A, and on heating, curve B of figure 23. The crystal structure remains face-centred cubic on going through the loop, but at the low temperature end a collapsed phase with 15% smaller volume appears (Schuch and Sturdivant 1950). Lock (1957 a), using susceptibility measurements, has shown that the hysteresis loop collapses on repeatedly cooling to 20°K. The current explanation is that on cooling through the loop the 4f electron goes into a 5d state. This fits in with evidence from the susceptibility measurements, electrical conductivity and change in volume. There is no proper explanation of why such a transition should exhibit hysteresis or of how the effect is suppressed by repeated coolings. At lower temperatures a large rounded peak is found in the specific heat curve at 12·5°K. The height of this anomaly depends on the number of coolings

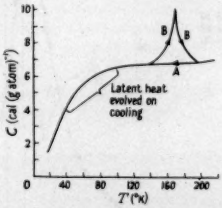


Figure 23. Hysteresis in the specific heat of cerium.

as would be expected from the hysteresis behaviour if it is due to the 4f electrons (Parkinson and Roberts 1957). At the same temperature the susceptibility (Lock 1957) goes through a maximum and becomes field dependent and the general behaviour is that which would be expected if the metal becomes anti-ferromagnetic at 12·5°K. An approximation to the 'anomalous entropy' under the specific heat curve can be found by subtracting the specific heat of lanthanum. The value so found is approximately 1·3, which is close to $R\ln 2$. This metal forms solid solutions with lanthanum, so on alloying magnetic dilution can be achieved. The Néel point moves to lower temperatures as the concentration of lanthanum increases (Lock and Roberts 1957), but the specific heat curve breaks into two humps with the Néel point corresponding to the lower one. Nevertheless, the entropy under both humps must be used to give a value per atom corresponding to that of pure cerium. The upper specific heat hump is probably associated with some sort of short-range ordering process.

Praseodymium has a very large distributed anomaly centred at about 50°K. The susceptibility is always field independent and goes uniformly to a constant value at the lowest temperatures. Parkinson, Simon and Spedding (1951) explained this anomaly and those in cerium and neodymium, using the idea of Russell–Saunders coupling to give the ground state of the trivalent ions in the metals. This state could be split by electric fields due to the adjacent atoms (Stark splitting) and the anomalies could then be produced by exciting the electrons through the pattern of split levels. If a cubic electric field is assumed the anomalous hump

in praseodymium can be given a reasonable quantitative explanation, but this is the only instance where this explanation works at all well. With cerium and neodymium the positions of the anomalies are about right for such an explanation, but their shapes are quite wrong. Incidentally, with cerium, if this explanation is used together with the idea of strong magnetic interactions, there should be a second anomaly between 40° and 100°K. This has not been looked for with the metal in a 'cyclic state' (i.e. repeatedly cooled to suppress hysteresis effects), though there is evidence of the beginning of such behaviour in the work of Parkinson, Simon and Spedding (1951).

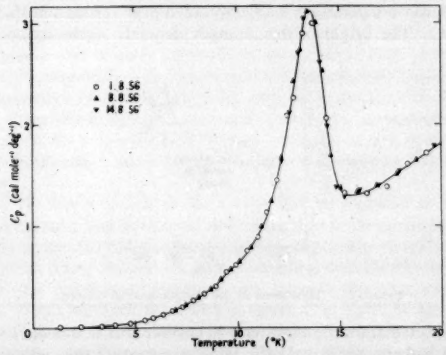


Figure 24. The specific heat of samarium. (*After* Roberts 1957 c.)

Neodymium is remarkable, showing two pronounced rounded peaks in the specific heat, one at 7·5°K and the other at 20°K. At 7·5°K the susceptibility reaches a maximum and below this temperature becomes field dependent, and the evidence points to the metal becoming anti-ferromagnetic. At 20°K the curve of inverse susceptibility shows a sharp change in gradient. This higher anomaly is not properly understood, but like those in the alloys of cerium and lanthanum, is probably due to a short-range ordering mechanism.

Element 61, promethium, does not exist naturally and sufficient quantities do not seem to have been produced artificially for the specific heat to be known yet. A behaviour similar to praseodymium could be expected. Samarium has five 4f electrons and its specific heat has only been measured below 20°K. It has a very large and sharp anomaly at 14·5°K (figure 24, Roberts 1957 c). The susceptibility (Lock 1957 a) also reaches a sharp maximum at the same temperature and points to anti-ferromagnetism. No specific heat measurements have been made on europium yet. A curve of atomic volume plotted against atomic number for these metals and the chemical properties show that the 5d electron in europium goes into the 4f shell, so making seven 4f electrons in all, which is a stable configuration. The original work of Klemm and Bommer (1937) suggests a

paramagnetic Curie point at about 15°K; behaviour similar to gadolinium can otherwise be expected.

The ferromagnetism of gadolinium, with a Curie temperature of 289°K, is probably the best known of any properties of the rare earth metals. It has a large λ-point anomaly at 291·8°K (Griffel, Skochdopole and Spedding 1954), but the specific heat measurements only go down to 15°K. The next five metals—terbium, dysprosium, holmium, erbium and thulium—in which n runs from 8 to 12, all show complex anti-ferromagnetic and ferromagnetic behaviour. Where measurements exist the expected peaks in the specific heat curves have been observed.

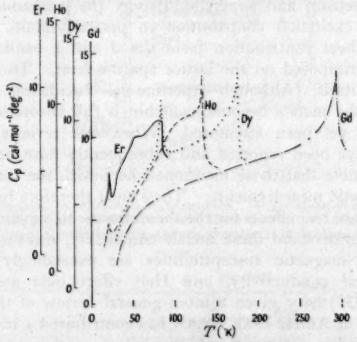


Figure 25. The specific heats of gadolinium, dysprosium, holmium and erbium. Note off-set heat capacity scales.

Terbium is paramagnetic down to about 230°K at which point it has a ferromagnetic Curie point; no specific heat data exist yet. Dysprosium is paramagnetic down to 175°K, where it becomes anti-ferromagnetic, and then at 85°K it becomes ferromagnetic (Elliott, Legvold and Spedding 1954). There is a sharp λ-point in the specific heat curve at 174°K and at 83·5°K a symmetrical peak (Griffel, Skochdopole and Spedding 1956). Holmium behaves in a similar way, having a Néel point at 132°K and a ferromagnetic Curie point about 20°K. There is a sharp λ-point at the Néel temperature and a symmetrical peak at 19·4°K. There are also two points of inflection between the two transitions. Erbium is also similar, but now the magnetic transitions are more smeared out. It is paramagnetic down to 80°K where there is a slight rise in susceptibility, which probably corresponds to a Néel point. The specific heat curve has a sharp rise at 84°K, but the anomaly has a flat top (figure 25), while at 19·9° there is a symmetrical peak corresponding to the onset of ferromagnetism. In addition, between these two points there is a third rounded maximum which does not correspond to any particular magnetic phenomenon. With dysprosium, holmium and erbium, the specific heat at the ferromagnetic Curie point shows hysteresis effects. Thulium is paramagnetic down to 51°K, where it appears to become anti-ferromagnetic,

and so far there is no evidence of ferromagnetism at lower temperatures; no specific heat measurements have yet been made. In ytterbium, like europium, the 5d electron goes into the 4f shell, but now to complete it; the metal does not therefore show any striking magnetic phenomena and it is not superconducting down to 1·2°K (Lock 1957 b). Likewise lutecium has a closed 4f shell and is not a superconductor down to 2·2°K. With these two metals no specific heat measurements at low temperatures have yet been reported.

From this summary it is evident that these metals are remarkable in that their behaviour includes very nearly every type of specific heat phenomenon that could be expected in a metal, from the co-operative phenomena of ferromagnetism and anti-ferromagnetism, and superconductivity (in lanthanum), to a Schottky type of electronic excitation contribution in praseodymium. These mask the additional specific heat contribution from the d and s bands which must also exist, and are superimposed on the lattice specific heat. To deal with them in full is a report in itself. Although experimental elucidation of their properties is going ahead as the metals become available, a full theoretical attack on their behaviour has not yet been attempted. When one reviews the specific heat 'bumps' which have been reported and subsequently have not been confirmed, it is heartening to note that those mentioned here with the rare-earth metals are supported by magnetic measurements. They must therefore be regarded as quite genuine, and there are real effects for the theoreticians to explain, however difficult this may be. To understand these metals completely, measurements other than specific heats and magnetic susceptibilities are exceedingly useful. Neutron diffraction, electrical conductivity, and Hall effect data are being collected. Spedding *et al.* (1957) have given a more general review of the work, primarily of their own group at Ames, Iowa, which has contributed a major portion of the experimental knowledge of these metals.

§6. Experimental Techniques

There is no point in cataloguing all the different experimental details which have been or could be used in low temperature calorimetry, but there are a number of general aspects which are worth discussing.

The usual techniques are based on the vacuum calorimeter of Nernst. The calorimeter is mounted inside an enclosure, normally called the shield, which is held at a constant temperature and which is highly evacuated; figure 26(*a*) and (*b*) shows two typical arrangements. At low temperatures the only thermal contact with the shield or any other surroundings should be through the electrical leads and other necessary connections to the calorimeter. In working at low temperatures, far more complete insulation is possible than in normal room temperature calorimetry.

The calorimeter must first be cooled to a temperature very near to that of the shield and the method of so doing will be discussed shortly. Then a measured quantity of heat can be put into the calorimeter electrically and the corresponding temperature rise observed. The drift in temperature before and after the heating interval must also be recorded (figure 27) and it is usual to run a whole sequence of points one after the other. Usually the heating is carried out by passing a known current i through a constantan heater of resistance R, mounted in the

calorimeter. In the liquid helium and hydrogen temperature ranges R for constantan varies by about 0·1% per degree and this variation has to be measured; in spite of it the current can be kept constant enough by using large fixed resistances in the heater circuit. Obviously the heat generated is $Q = Ri^2t/J$ where t is the time interval and J is Joule's constant. The resistance R and heating currents can be measured with high precision using normal d.c. potentiometer techniques and the largest source of error in determining Q lies in the measurement of t.

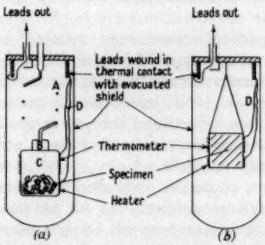


Figure 26. Typical schematic experimental arrangements: (*a*) specimen in calorimeter; (*b*) specimen regular shape with directly wound heater and thermometer. A, condensation tube; B, condensation chamber; C, calorimeter specimen chamber; D, leads to calorimeter.

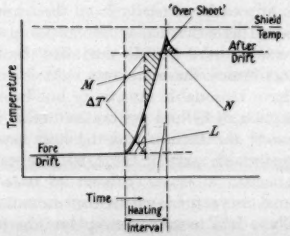


Figure 27. Typical temperature–time curve for one point.

Reliable mechanical devices for simultaneously operating both the electrical switch gear and a calibrated stop watch are easily devised. There are, however, always uncertainties over the precise switching time and errors in reading a stop watch. Consequently, the heating interval has to be not less than 30 seconds and may be as long as 180 seconds. The precision of switching and timing can be vastly improved by using electronic controls which count in whole numbers of seconds and only switch at the end of a completed second.

The calorimeter temperature is usually measured with suitable resistance thermometers. Constantan has been used in the helium–hydrogen temperature range but is very insensitive particularly below 5°K and often has irregularities in its resistance–temperature characteristic between 7° and 8°K. For the helium

temperature range cold drawn phosphor–bronze and leaded-brass wires have high sensitivity (cf. Parkinson and Roberts 1955). Both depend on the effects of inclusions of superconducting lead on their resistance–temperature characteristics. In consequence they are sensitive to external magnetic fields and dependent on the measuring current.

In the last few years carbon resistors (Allen and Bradley) have been introduced very successfully as thermometers by Clement and Quinnell (1952). They have a low heat capacity and are small (tenth watt size). Their resistance–temperature law is of the approximate form $\log R + k/\log R = A + B/T$; but for the precision requirements of calorimetry detailed calibrations of individual thermometers are always necessary. The chief difficulty is to choose a resistance which does not become inconveniently high at the lowest temperatures of operation. Clement and Quinnell (1952) have found a method of making this choice from room temperature data. Provided the power dissipation in the resistors is kept low, they give reproducible results and have a well-defined variation of resistance with applied magnetic field.

At higher temperatures platinum resistance thermometers are usual, for they are reliable and have high sensitivity. As far as measuring the resistance is concerned, there is little to choose between using a potentiometer or Wheatstone bridge. The latter can save one lead per thermometer to the calorimeter and so cut unwanted thermal leaks. On the other hand, the lead resistances have to be carefully balanced to give proper compensation which is unnecessary with a potentiometer method. With either method the limit to precision is usually set by stray thermo-electric effects, particularly from the leads which inevitably have junctions while also having high thermal gradients along their length.

Thermocouples have been used but do not offer the high sensitivities which can be achieved with resistance thermometers. At lower temperatures helium gas thermometers also have reasonable sensitivity but below 4°K become troublesome owing to the adsorption of helium gas on the inside walls of the bulb. The latent heat of desorption of the helium and the heat capacity of the gas in the bulb can become an appreciable part of the total heat capacity of the specimen and calorimeter system.

The temperature–time curve for a particular measurement can be recorded manually or automatically. It is usual to extrapolate the fore and after drifts into the heating interval and to take the temperature rise ΔT as being the interval between the extrapolated lines at the centre of the heating period (figure 27). Strictly, ΔT should be measured at the point where the shaded areas $L+N$ equal M. The fore and after drift curves are parts of exponential curves but with good thermal insulation should not be observably different from straight lines. Parkinson and Quarrington (1954 a) have given an analysis of the errors which can occur with this method of calculation when one is forced to use a separate calorimeter enclosing the specimen. This is so with the highly oxidizable alkali metals, for example, or with collections of crystals of the element under investigation, and under these conditions the thermal contact between the specimen and calorimeter is usually achieved using a small trace of helium exchange gas. Suppose the rate at which heat can be conducted between the specimen and calorimeter is P joule deg^{-1} sec^{-1}, and between the calorimeter and surroundings

is p joule deg^{-1} sec^{-1}. The normal method of measuring and calculation will then give rise to a systematic error and if C_o is the observed heat capacity of the specimen the true heat capacity is $C_t = C_o/\{1+C_o p/P(C_o+C_1)\}$ provided p is small compared with P and where C_1 is the heat capacity of the calorimeter. Errors of this sort can creep in and remain undetected when helium exchange gas is used to cool the calorimeter to the enclosure temperature. These effects can be reduced to negligible proportions by providing adequate pumping arrangements to the shield vacuum space (e.g. Corak, Garfunkel, Satterthwaite and Wexler 1955).

Where the metal to be investigated is more tractable, as for example copper, it can be machined to a suitable cylindrical shape and the thermometer and heater wound directly on to it (figure 26(*b*)). Thermal contact can be made good by using a suitable varnish such as damarda lacquer or araldite, to hold the windings in place. Windings can be put on to a calorimeter in a similar way. It is then necessary in either instance to show that the windings are in good contact with the specimen or calorimeter, which can be done by observing the thermometer response to changes in the heater current. There are a variety of ways in which calorimeters and specimens can be mounted and the form of the temperature–time curve recorded will depend on the geometry used and the quality of the various thermal contacts. The thermometer may 'overshoot', as shown in figure 27, owing to the proximity of the heater. This is not necessarily objectionable, though it is undesirable, neither is the lack of overshoot necessarily an indication of good conditions for accurate calorimetry. The proper conditions for working can only be found by considering the equivalent thermal circuit of the calorimeter and shield system in the manner of Parkinson and Quarrington (1954 a) and also Keesom and Kok (1934). It is then necessary to find by experiment the relevant 'thermal transfer coefficients' such as p and P referred to above.

It is usual to cool the calorimeter either by using exchange gas or by condensing a small quantity of helium liquid into a chamber in the calorimeter. The condensation method has two advantages: the surrounding vacuum space is not disturbed and contaminated by residual traces of exchange gas desorbed in the course of the experiment, and the condensed gas can be very conveniently used for calibrating or checking the thermometers. The condensed helium liquid is removed from the calorimeter by pumping. There are, however, two disadvantages. First, the condensation chamber may retain considerable quantities of adsorbed helium on its inner surface for which the latent heat of desorption can be appreciable; and second, it is necessary to connect the calorimeter to the shield by a thin walled tube which, even though made of cupronickel, can provide a larger thermal connection than all the leads put together.

The use of helium exchange gas in the shield overcomes these two disadvantages, but at the lowest temperatures desorption from the calorimeter walls can be troublesome unless very good pumping facilities are available to bring the shield vacuum to the order of 10^{-6} mm at the lowest temperatures (cf. Corak *et al.* 1954). Experiments must be conducted with either system to ensure that interference from either internally or externally adsorbed helium is negligible. The use of helium exchange gas within the calorimeter itself to make thermal contact with

the specimen can also be objectionable. A filling pressure of only about 0·1 mm of mercury at room temperature is quite enough to produce interference in the temperature range between 1° and 2°K with specimens of low heat capacity (cf. Roberts 1957 a).

A solution to these difficulties would be to use a thermal switch between the calorimeter and shield, which could be mechanical. The rate at which heat can be driven across a mechanical contact is strictly limited by the real area of contact. Normally there are only three points of contact and considerable pressure is needed to give a reasonable thermal transfer rate. In consequence the time taken in the initial cooling from say 10 or 20°K to the operating temperature may be quite long. Ramanathan and Srinivasan (1955) have devised a method in which the sample, wound with heater and leads, has a polished bottom which can be pressed against a polished surface on the shield itself. In this way cooling could be effected successfully from liquid air temperatures, and difficulties from desorbed gas completely eliminated. Whatever switch is used it must connect the calorimeter to a helium bath for thermometer calibration purposes, and therefore it must be capable of conducting heat at a rate far greater than the dissipation in the thermometers, for very small temperature differences across it. For similar reasons all leads must come into good thermal contact with the bath before going to the calorimeter. Usually the leads are anchored thermally to the shield by sticking them in place with varnish (e.g. araldite). Lengths of about 15 to 20 cm in contact with the shield are the minimum for effective anchoring under normal conditions.

The heater leads can be troublesome for their resistance must be only a fraction of that of the heater, which tends to imply thick short leads, whereas good thermal insulation demands thin long leads. A compromise has to be achieved using as high a heater resistance as possible. In the helium range a good solution is to use thin constantan wires coated with a thin layer of lead. The lead is a superconductor in the operating range, so the leads have negligible electrical resistance while their small diameter and long length provided a large thermal resistance. The total heat leak between the shield and calorimeter need not be greater than 0·1 ergs per second for 1°K temperature difference.

There is considerable interest in carrying on calorimetric experiments on metals at temperatures below 1°K. Rayne has used a conventional technique in which the specimen and salt used for demagnetizing were thermally bound together throughout the experiment. The specimen was soldered to a thick copper wire which was soldered in turn to a set of copper vanes embedded in the salt. The heat capacity of the metal diminishes at least linearly with temperature towards zero, whereas that of the paramagnetic salt is increasing towards the peak of a specific heat anomaly associated with the electronic energy levels which make demagnetization possible. In this way at the lowest temperatures the heat capacity of the specimen rapidly becomes a small fraction of the total which is measured. Samoilov (1952) used rather long and very thin copper leads to separate his specimen and salt. They had enough thermal conductivity to enable the specimen to cool to the salt temperature in about 60 minutes, but on the other hand, sufficient thermal resistance for the specimen only to be heated during individual heating intervals.

In the construction of a calorimeter or in putting windings directly on to a specimen, varnishes and solders are necessary and quite small quantities of them can have objectionably high heat capacities in the helium range (Keesom and Pearlman 1952). The quantities used must be reduced to the essential minimum and be known accurately. Some useful data on the heat capacity of araldite varnish and Woods metal have been given by Parkinson and Quarrington (1954 b).

Howling, Mendoza and Zimmerman (1955) have described an indirect method of deducing the specific heat of a metal which is suitable for very low temperature operation. By periodically varying the temperature at the end of a specimen in the form of a small rod, the thermal diffusivity can be found. The same experimental set up can also be used to find the thermal conductivity of the specimen and hence the specific heat can be deduced.

In low temperature calorimetry the secondary standards used for thermometer calibrations are the vapour pressures of liquid helium and hydrogen. New tables of the vapour pressure curves of helium have been produced, based on the work of Van Dijk. Evidence for the need of revision was demonstrated in accurate calorimetric work, such as that of Corak, Garfunkel, Satterthwaite and Wexler (1955) on the noble metals. Van Dijk and Durieux (1957) have published an account of the establishment of the helium temperature scales and the difficulties in so doing.

§ 7. Conclusion

As was remarked in the introduction, a discussion of the specific heats of metals covers a very wide field, and although the scope of this paper is limited, a number of unresolved problems have been touched upon. Their solution will involve the study of many other parameters, for the specific heat is too insensitive to much of the 'finer structure of matter'. Thus there is adequate evidence to show that the effect of free electrons on the lattice specific heat must be taken into account in any complete theory. This implies an associated study of the elastic constants of the metal.

With the conduction electrons in the monovalent metals, the real interest lies in the shape and position of the Fermi surface. Although more specific heat data taken below 1°K are badly needed for the alkali metals, this problem may be resolved by other means as indicated by Pippard (1955, 1957) and Heine (1957 a, b, c). Perhaps one of the most interesting experiments would be a study of differences between the lattice and free electronic specific heats and other parameters in the isotopes lithium 6 and 7.

With the transition metals and superconductors the electronic specific heats are necessary basic data but the detailed description of the electronic structure will result from other studies, such as the microwave absorption in superconductors. With the rare-earth metals, also, the specific heats give indispensable information concerning the general nature and entropies involved in their anomalies; with this information in hand, susceptibility measurements are more revealing.

These final remarks simply emphasize the necessity of an 'all round' attack on any physical problem. However, from what has been said it is evident that a great deal more reliable information on the detailed behaviour of metals at low temperatures is required.

References

Bardeen, J., 1936, *Phys. Rev.*, **56**, 1698 ; 1955, *Ibid.*, **97**, 1724.

Bardeen, J., Cooper, L. N., and Schrieffer, J. R., 1957, *Phys. Rev.*, **108**, 1175.

Barrett, C. S., 1956, *Acta Cryst., Camb.*, **9**, 671.

Barron, T. H., and Morrison, J. A., 1957, *Canad. J. Phys.*, **35**, 799.

Bauer, E., 1953, *Phys. Rev.*, **92**, 58.

Berman, A., Zemansky, M. W., and Boorse, H. A., 1955, *Conférence de Physique des Basses Températures, Paris* (Paris : Inst. Int. du Froid), p. 499.

Bhatia, A. B., 1955, *Phys. Rev.*, **97**, 363.

Bhatia, A. B., and Horton, G. K., 1955, *Phys. Rev.*, **98**, 1715.

Bijl, D., 1957, *Progress in Low Temperature Physics*, Vol. II (Ed. C. J. Gorter) (Amsterdam : North-Holland), p. 395.

Biondi, M. A., Garfunkel, M. P., and McCoubrey, A. O., 1956, *Phys. Rev.*, **101**, 1427.

Blackman, M., 1935 a, *Proc. Roy. Soc.* A, **148**, 365, 385 ; 1935 b, *Ibid.*, **149**, 117, 126 ; 1937 a, *Ibid.*, **159**, 416 ; 1937 b, *Proc. Camb. Phil. Soc.*, **33**, 94 ; 1941, *Rep. Progr. Phys.*, **8**, 11 (London : Physical Society) ; 1951, *Proc. Phys. Soc.* A, **64**, 681 ; 1956, *Handbuck der Physik*, Vol. 7, Pt. 1 (Berlin : Springer), p. 325.

Blewins, G. S., Gordy, W., and Fairbanks, W. M., 1955, *Phys. Rev.*, **100**, 1215.

Bohm, D., and Pines, D., 1951, *Phys. Rev.*, **82**, 625 ; 1952, *Ibid.*, **85**, 338 ; 1953, *Ibid.*, **92**, 609.

Born, M., 1923, *Atom Theorie der Festen Zustanden* (Leipzig : Teubner).

Born, M., and Von Karman, Th., 1912, *Phys. Z.*, **13**, 219.

Brillouin, L., 1946, *Wave Propagation in Periodic Structures* (New York : McGraw Hill).

Buckingham, M. J., 1951, *Nature, Lond.*, **168**, 281.

Buckingham, M. J., and Schafroth, M. R., 1954, *Proc. Phys. Soc.* A, **67**, 828.

Chambers, R. G., 1956, *Canad. J. Phys.*, **34**, 1395.

Clement, J. R., and Quinnell, E. H., 1952, *Rev. Sci. Instrum.*, **23**, 213.

Corak, W. S., Garfunkel, M. P., Satterthwaite, C. B., and Wexler, A., 1955, *Phys. Rev.*, **98**, 1699.

Corak, W. S., Goodman, B. B., Satterthwaite, C. B., and Wexler, A., 1954, *Phys. Rev.*, **96**, 1442.

Cristescu, S., and Simon, F. E., 1934, *Z. Phys. Chem.*, **25B**, 273.

Daunt, J. G., 1955, *Progress in Low Temperature Physics*, Vol. I (Ed. C. J. Gorter) (Amsterdam : North-Holland), p. 202.

Daunt, J. G., and Horowitz, M., 1953, *Phys. Rev.*, **91**, 1099.

Daunt, J. G., Horseman, A., and Mendelssohn, K., 1939, *Phil. Mag.*, **27**, 754.

Daunt, J. G., and Mendelssohn, K., 1937, *Proc. Roy. Soc.* A, **160**, 127.

Daunt, J. G., and Silvidi, A. A., 1950, *Phys. Rev.*, **77**, 125.

Debye, P., 1912, *Ann. Phys., Lpz.*, **39**, 789.

Dolecek, R. L., 1955, *Conférence de Physique des Basses Températures, Paris* (Paris : Inst. Int. du Froid), p. 300.

Einstein, A., 1907, *Ann. Phys., Lpz.*, **22**, 180, 800.

Eisenstein, J., 1954, *Rev. Mod. Phys.*, **26**, 277.

Elliott, J. F., Legvold, S., and Spedding, F. H., 1954, *Phys. Rev.*, **94**, 1143.

Estermann, I., Friedberg, S. A., and Goldman, J. E., 1952, *Phys. Rev.*, **87**, 582.

Estermann, I., and Weertman, S. A., 1952, *J. Chem. Phys.*, **20**, 972.

Eucken, A., 1913, *Verh. dtsch. Phys. Ges.*, **15**, 571.

Fine, P. C., 1939, *Phys. Rev.*, **56**, 355.

Fletcher, G. C., 1952, *Proc. Phys. Soc.* A, **65**, 192.

Fletcher, G. C., and Wohlfarth, E. P., 1951, *Phil. Mag.*, **42**, 104.

Friedberg, S. A., 1954, Carnegie Institute of Technology, Dept. Phys., Final Report on Contract, N6ORI-47-T.O.3.

Fröhlich, H., 1950, *Phys. Rev.*, **79**, 845 ; 1954, *Proc. Roy. Soc.* A, **223**, 296.

Fuchs, K., 1936, *Proc. Roy. Soc.* A, **153**, 622 ; 1936, *Ibid.*, **157**, 444.

Glover, R. E., and Tinkham, M., 1957, *Phys. Rev.*, **108**, 243.

GOODMAN, B. B., 1955, *Conférence de Physique des Basses Températures, Paris* (Paris : Inst. Int. du Froid), p. 506.
GRAYSON-SMITH, H., and STANLEY, J. P., 1950, *J. Chem. Phys.*, **18**, 136.
GRIFFEL, M., SKOCHDOPOLE, R. E., and SPEDDING, F. H., 1954, *Phys. Rev.*, **93**, 675 ; 1956, *J. Chem. Phys.*, **25**, 75.
GRUNEISEN, E., and GOENS, E., 1924, *Z. Phys.*, **26**, 250.
HEINE, V., 1957 a, *Proc. Roy. Soc.* A, **240**, 340 ; 1957 b, *Ibid.*, **240**, 354 ; 1957 c, *Ibid.*, **240**, 361.
HERRING, C., and HILL, A. G., 1940, *Phys. Rev.*, **58**, 132.
HILL, R. W., and PARKINSON, D. H., 1952, *Phil. Mag.*, **43**, 309.
HILL, R. W., and SMITH, P. L., 1953, *Phil. Mag.*, **44**, 636.
HOARE, F. E., 1955, *Conférence de Physique des Basses Températures, Paris* (Paris : Inst. Int. du Froid), p. 304.
HOUSTON, W. V., 1948, *Rev. Mod. Phys.*, **20**, 161.
HOWLING, D. H., MENDOZA, E., and ZIMMERMAN, J. E., 1955, *Proc. Roy. Soc.* A, **229**, 86.
HSIEH, Y-C., 1954, *J. Chem. Phys.*, **22**, 306.
IONA, M., 1941, *Phys. Rev.*, **60**, 823.
JONES, H., 1937, *Proc. Phys. Soc.*, **49**, 250 ; 1955, *Proc. Phys. Soc.* A, **68**, 1191 ; 1957, *Proc. Roy. Soc.* A, **240**, 321.
JONES, H., and MOTT, N. F., 1937, *Proc. Roy. Soc.* A, **162**, 49.
KEESOM, P. H., and PEARLMAN, N., 1952, *Phys. Rev.*, **88**, 398, 140 ; 1954, *Ibid.*, **96**, 897 ; 1955, *Ibid.*, **98**, 548 ; 1956, *Handbuch der Physik*, Vol. 14 (Berlin : Springer), p. 282.
KEESOM, W. H., and CLARK, C. W., 1935, *Physica*, **2**, 513.
KEESOM, W. H., and KOK, J. A., 1934, *Physica*, **1**, 770.
KEESOM, W. H., and KURRELMEYER, B., 1940, **10**, 1003.
KEESOM, W. H., and VAN DER ENDE, J. N., 1932, *Proc. Acad. Sci. Amst.*, **35**, 143.
KELLERMANN, E. W., 1941, *Proc. Roy. Soc.* A, **178**, 17.
KLEMENS, P. G., 1954, *Aust. J. Phys.*, **7**, 64.
KLEMM, W., and BOMMER, H., 1937, *Z. anorg. Chem.*, **231**, 138.
KOPPE, H., 1947, *Z. Naturforsch.*, **2a**, 185 ; **3a**, 65.
KOSTER, G. F., 1955, *Phys. Rev.*, **98**, 901.
KREISSMAN, C. J., 1953, *Rev. Mod. Phys.*, **25**, 122.
KREISSMAN, C. J., and MCGUIRE, T. R., 1955, *Phys. Rev.*, **98**, 936.
KROLL, W., 1952, *Progr. Theor. Phys., Japan*, **8**, 457.
KRUTTER, H., 1935, *Phys. Rev.*, **48**, 664.
LANDSBERG, P. T., 1948, *Proc. Phys. Soc.* A, **62**, 806.
LANGE, F., 1924, *Z. Phys. Chem.*, **110**, 343.
DE LAUNAY, J., 1953, *J. Chem. Phys.*, **21**, 1975 ; 1954, *Ibid.*, **22**, 1676 ; 1956 a, *Ibid.*, **24**, 1071 ; 1956 b, *Solid State Physics*, **2**, 220.
LEE-WHITING, G. E., 1952, *Proc. Roy. Soc.* A, **212**, 362.
LEIBFRIED, G., and BRENNIG, W., 1953, *Z. Phys.*, **134**, 451.
LEIGH, R. S., 1951, *Phil. Mag.*, **42**, 139.
LEIGHTON, R. B., 1948, *Rev. Mod. Phys.*, **20**, 165.
LLYDANE, R. H., and HERZFELD, K. F., 1938, *Phys. Rev.*, **54**, 846.
LOCK, J. M., 1957 a, *Proc. Phys. Soc.* B, **70**, 566 ; 1957 b, *Ibid.*, **70**, 467.
LOCK, J. M., PIPPARD, A. B., and SHOENBERG, D., 1951, *Proc. Camb. Phil. Soc.*, **47**, 811.
LOCK, J. M., and ROBERTS, L. M., 1957, *Phil. Mag.* (8), **2**, 811.
MADELUNG, E., 1909, *Nachr. Ges. Wiss. Göttingen Math. Phys. Kl.*, **1**, 100.
MANNING, M. F., 1943, *Phys. Rev.*, **63**, 190.
MANNING, M. F., and CHODOROW, M., 1939, *Phys. Rev.*, **56**, 787.
MANNING, M. F., and KRUTTER, H. M., 1937, *Phys. Rev.*, **51**, 761.
MATYAS, Z., 1948, *Phil. Mag.*, **39**, 429.
MAXWELL, E., 1952, *Phys. Rev.*, **86**, 235.
MAXWELL, E., and LUTES, O. S., 1954, *Phys. Rev.*, **95**, 333.
MONTROLL, E., 1942, *J. Chem. Phys.*, **10**, 218 ; 1943, *Ibid.*, **11**, 481.
MONTROLL, E., and PEASLEE, D., 1944, *J. Chem. Phys.*, **10**, 218.

MOTT, N. F., 1935, *Proc. Phys. Soc.*, **47**, 571.
MOTT, N. F., and STEVENS, K. W. H., 1957, *Phil. Mag.* (8), **2**, 1364.
PARKINSON, D. H., and QUARRINGTON, J. E., 1954 a, *Proc. Phys. Soc.* A, **57**, 569 ; 1954 b, *Brit. J. Appl. Phys.*, **5**, 219 ; 1955, *Proc. Phys. Soc.* A, **58**, 762.
PARKINSON, D. H., and ROBERTS, L. M., 1955, *Proc. Phys. Soc.* B, **58**, 386 ; 1957, *Ibid.*, **70**, 471.
PARKINSON, D. H., SIMON, F. E., and SPEDDING, F. H., 1951, *Proc. Roy. Soc.* A, **207**, 137.
PICKARD, G. L., and SIMON, F. E., 1948, *Proc. Phys. Soc.*, **61**, 1.
PINES, D., 1955, *Solid State Physics*, **1**, 367.
PIPPARD, A. B., 1955, *Proc. 10th Solvay Congress* (1954), Brussels ; 1957, *Proc. Roy. Soc.* A, **250**, 325.
RAMAN, C. V., 1941, *Proc. Indian Acad. Sci.* A, **14**, 459.
RAMANATHAN, K. G., and SRINIVASAN, T. M., 1955, *Phil. Mag.*, **46**, 338.
RAYNE, J., 1954, *Phys. Rev.*, **95**, 1428 ; 1956 a, *Aust. J. Phys.*, **9**, 189 ; 1956 b, *Proc. Phys. Soc.* B, **69**, 482.
RAYNOR, G. V., 1952, *Rep. Progr. Phys.*, Vol. XV, p. 173 (London : Physical Society).
REITZ, J. R., 1955, *Solid State Physics*, **1**, 1.
ROBERTS, L. M., 1957 a, *Proc. Phys. Soc.* B, **70**, 738 ; 1957 b, *Ibid.*, **70**, 744 ; 1957 c, *Ibid.*, **70**, 434.
SAMOILOV, B. N., 1952, *Dokl. Akad. Nauk. S.S.S.R.*, **86**, 281.
SCHROEDINGER, E., 1926, *Handbuch der Physik*, Vol. 10 (Berlin : Springer).
SCHUCH, A. F., and STURDIVANT, J. H., 1950, *J. Chem. Phys.*, **18**, 145.
SERIN, B., REYNOLDS, C. A., and LOHMAN, A., 1952, *Phys. Rev.*, **86**, 162.
SHOMATE, Y. H., 1945, *J. Chem. Phys.*, **13**, 326.
SHULL, C. G., and WILKINSON, M. K., 1953, *Rev. Mod. Phys.*, **25**, 100.
SIMON, F. E., 1930, *Ergbn. exakt. Naturw.*, **9**, 222.
SKINNER, H. W. B., 1939, *Rep. Progr. Phys.*, Vol. V, p. 257 (London : Physical Society) ; 1940, *Phil. Trans. Roy. Soc.* A, **239**, 12.
SLATER, J. C., 1936, *Phys. Rev.*, **49**, 537.
SLATER, J. C., and KOSTER, G. F., 1954, *Phys. Rev.*, **94**, 1498.
SLUTSKY, L. J., and GARLAND, C. W., 1957, *J. Chem. Phys.*, **26**, 787.
SMITH, H. J., 1947, *Phil. Trans.* A, **188**, 179.
SMITH, P. L., 1955, *Conférence de Physique des Basses Températures, Paris* (Paris : Inst. Int. du Froid), p. 281.
SNIDER, J. L., and NICOL, J., 1957, *Phys. Rev.*, **105**, 1242.
SOMMERFELD, A., 1928, *Z. Phys.*, **47**, 1.
SOMMERFELD, A., and BETHE, H., 1934, *Handbuch der Physik*, Vol. 24/2 (Berlin : Springer).
SPEDDING, F. H., LEGVOLD, S., DRANE, A. H., and JENNINGS, L. D., 1957, *Progress in Low Temperature Physics*, Vol. II (Amsterdam : North-Holland), p. 368.
STONER, E. C., 1936, *Phil. Mag.* (7), **21**, 145.
STOUT, J. W., and GUTTMAN, L., 1952, *Phys. Rev.*, **88**, 703.
SUTHERLAND, W., 1910, *Phil. Mag.* (6), **20**, 657.
TINKHAM, M., 1956, *Phys. Rev.*, **104**, 845.
VAN DIJK, H., and DURIEUX, M., 1957, *Progress in Low Temperature Physics*, Vol. II (Amsterdam : North-Holland), p. 431.
WAITE, T. R., CRAIG, R. S., and WALLACE, W. E., 1956, *Phys. Rev.*, **104**, 1240.
WEXLER, A., and CORAK, W. S., 1952, *Phys. Rev.*, **85**, 85.
WIGNER, E., 1947, *Trans. Faraday Soc.*, **34**, 678.
WOHLFARTH, E. P., 1949, *Proc. Roy. Soc.* A, **195**, 434 ; 1950, *Phil. Mag.*, **41**, 534.
WOLCOTT, N. M., 1955, *Conférence de Physique des Basses Températures, Paris* (Paris : Inst. Int. du Froid), p. 286.
WRIGHT, S. J., 1930, *Proc. Roy. Soc.* A, **126**, 613.
YATES, B., and HOARE, F. E., 1957, *Proc. Roy. Soc.* A, **240**, 42.
ZWIKKER, C., 1928, *Z. Phys.*, **52**, 668.

MESONIC ATOMS

By D. WEST

Atomic Energy Research Establishment, Harwell, Berks.

CONTENTS

	PAGE
§ 1. Introduction	272
§ 2. Energy levels: theoretical	273
2.1. Energy levels of the mesonic atom in a Coulomb potential	273
2.2. Nuclear size effect	274
2.3. Vacuum polarization	276
2.4. Other perturbations	277
2.5. Fine structure	277
§ 3. Cascade of transitions to the ground state	278
§ 4. Measurements of the energies of mesonic x-rays	280
4.1. Experimental techniques	280
4.2. Measurements of low energy x-rays and the determination of the meson masses	282
4.3. Measurements of high energy μ-mesonic x-rays and the determination of nuclear radii	284
4.4. π-mesonic x-rays and the nuclear force shift	287
§ 5. Experimental determination of the yields of K and L series x-rays	292
5.1. X-ray yields from μ mesonic atoms	292
5.2. X-ray yields from π mesonic atoms	295
5.3. Relative yields of higher transitions	297
§ 6. Nuclear capture from the 1s level	298
6.1. μ mesons: lifetime measurements	298
6.2. π mesons: level width measurements	299
§ 7. Relative probability of the capture of a meson in the constituents of a chemical compound	302
§ 8. Mesonic molecules and the catalysis of nuclear reactions between hydrogen isotopes	304
§ 9. Mesonic atoms formed with K mesons and hyperons	307
Acknowledgments	308
Appendix. The overlap of Schroedinger wave functions with the nucleus	308
References	310

Abstract. The properties of mesonic atoms are described. The information which can be deduced about the properties of the π and μ mesons and their interactions with the nucleus is summarized. Experimental results on the nuclear size, meson masses and the π meson–nucleon interaction which have been obtained from a study of the energy levels are discussed. The experimentally determined yields of K and L series x-rays also give information on the processes of de-excitation of mesonic atoms. In particular there is evidence of collision de-excitation which is related to meson transfer in compounds and the formation of mesonic molecules which is observed in liquid hydrogen. Other properties of mesonic atoms at present beyond the reach of experimental investigation are briefly mentioned.

§ 1. Introduction

A mesonic atom consists of a normal atom with one of its electrons replaced by a negative meson. The existence of such atoms was first postulated in discussing the experimental result of Conversi *et al.* (1947) that stopped negative μ mesons were observed to decay in light but not in heavy elements. Fermi and Teller (1947) calculated that the time required for a slow (2000 ev) μ meson to reach its K level in graphite was only about 10^{-13} sec and so the radioactive decay observed (mean life $2{\cdot}2 \times 10^{-6}$ sec) would be expected to occur when the μ meson was bound in its K orbit. In heavier elements capture of the meson into the nucleus competes with radioactive decay. Wheeler (1947) calculated that for a meson bound in a 1s level the probability of nuclear capture would vary with atomic number as Z^4. Ticho (1948) measured the apparent lifetime of negative mesons stopped in elements ranging in atomic number from 8 to 16. The probability of nuclear capture deduced from the apparent lifetime was shown to vary with atomic number as $Z^{3{\cdot}7 \pm 0{\cdot}85}$. This observation constituted the first direct experimental evidence for mesonic atom formation. Further evidence was provided by the observation in photographic emulsions (Cosyns *et al.* 1949) of Auger electrons at the end of μ-meson tracks, and a similar observation (Menon *et al.* 1950) with π mesons. X-rays corresponding to transitions between the Bohr levels of the μ meson during its cascade down to the K shell were probably first observed by Chang (1949). However his measurements did not exclude the possibility that a large part of the radiation observed came from the nucleus following capture of the μ meson. More conclusive evidence of x-ray emission was obtained by Camac *et al.* (1952) who studied the radiation, emitted by π^- mesons stopped in graphite, by means of a scintillation counter. They obtained a peak in the vicinity of 100 kev, the expected energy of the K_α transition from carbon.

Subsequent interest in mesonic atoms has derived from the fact (Wheeler 1949) that the orbit radii are *smaller* than for an electron in the ratio (mass of meson) : (mass of electron), i.e. 207 for a μ meson and 273 for a π meson ; the binding energies of the levels are larger than for an electron by the same factor. One consequence of the small orbit size is that the wave functions of the levels of low principal quantum number lie completely within the electron shells. Consequently the meson's energy levels are hydrogen-like to a very good approximation and can therefore be calculated with great accuracy. However the small orbit size also means that an appreciable part of the wave function of the lowest level lies within the nucleus at intermediate atomic numbers. The nuclear radius is already equal to the unperturbed Bohr radius for a μ meson in its K shell at $Z = 45$. The overlapping of the wave function with the nucleus reduces the binding energy of the levels by an appreciable amount since the full Coulomb field is only effective outside the nucleus. With μ mesons the direct interaction of the meson with nuclear matter is completely negligible compared with the Coulomb force and exact calculation of the energy levels is still possible. The energy levels are sensitive to the value of the nuclear radius and measurements of the $2p \rightarrow 1s$ transition energies in intermediate and heavy elements by Fitch and Rainwater (1953) gave a very precise value for the nuclear radius.

The nuclear size effect is very small for the higher levels especially in light elements, and measurements of the L and M transitions have been used to determine

accurate values of the π- and μ-meson masses. μ-meson transitions were measured for this purpose by Koslov *et al.* (1954) and π-meson transitions by Stearns, Stearns, DeBenedetti and Leipuner (1954 a).

With π mesons the overlapping of the wave function with the nucleus leads to a very rapid absorption of the meson due to the strong interaction between π mesons and nuclei. The main effect of this is that the meson is usually captured into the nucleus before it can reach the lowest levels. Only in the lightest elements up to $Z = 9$ or so can the K x-rays be observed. Other consequences of the strong interaction of π mesons with nuclei are an energy shift of the 1s level which is observable even in the very light elements (Stearns, Stearns, DeBenedetti and Leipuner 1954 b), and an appreciable broadening of the 1s level (West and Bradley 1957).

Much theoretical work has been done on the finer details of mesonic x-ray spectra but, with present intensities of machine produced mesons, one is limited to using measuring instruments of relatively poor resolution. So far only scintillation counters and proportional counters can be used and a great deal of useful information is still out of reach.

§ 2. Energy Levels: Theoretical

2.1. *Energy Levels of the Mesonic Atom in a Coulomb Potential*

The main features of the energy levels for π and μ mesonic atoms can be seen most readily from the Bohr equations (1), (2) and (3).

$$E_n = -\bar{\mu}c^2\frac{(Z\alpha)^2}{2n^2} \qquad \ldots\ldots(1)$$

$$r_n = \frac{\hbar^2}{\bar{\mu}e^2}\frac{n^2}{Z} \qquad \ldots\ldots(2)$$

$$v_n = \alpha c\frac{Z}{n} \qquad \ldots\ldots(3)$$

where E_n is the energy of the level of principal quantum number n, α is the fine structure constant (1/137·04), $\bar{\mu}$ is the reduced mass of the meson nucleus system ($\bar{\mu} = \mu/(1+\mu/A)$), r_n is the radius of the nth orbit and v_n is the classical velocity in the orbit.

The dependence of E_n and r_n on the meson's reduced mass is evident together with the fact that relativistic effects will only be as important for mesons as for electrons since v_n is independent of $\bar{\mu}$. The exact energies of the levels in the absence of nuclear size effects can be obtained, in the case of π mesons, from solutions of the relativistic Schroedinger (Klein–Gordon) equation which applies to particles without spin. The exact solution of the Klein–Gordon equation (see, for example, Marshak 1952) can be written in the form given in equation (4),

$$E_{n,l} = \frac{-\bar{\mu}c^2}{2n^2}(Z\alpha)^2\left\{1+\frac{(Z\alpha)^2}{n^2}\left(\frac{n}{l+\frac{1}{2}}-\frac{3}{4}\right)-\ldots\right\} \qquad \ldots\ldots(4)$$

where terms in $(Z\alpha)^6$ and higher have been neglected and $Z < 137/2$. There is the usual small relativistic splitting of the states of different l for a given n, the s state being the most tightly bound.

For μ mesons, the energy levels are given by the solution of the Dirac equation for spin $\frac{1}{2}$ particles† (equation (5))

$$E_{n,j} = \frac{-\tilde{\mu}c^2}{2n^2}(Z\alpha)^2\left\{1+\frac{(Z\alpha)^2}{n^2}\left(\frac{n}{j+\frac{1}{2}}-\frac{3}{4}\right)-\ldots\right\} \qquad \ldots\ldots(5)$$

which is of the same form as equation (4) with j replacing l. The spin splitting between the levels $j = l+\frac{1}{2}$ and $j = l-\frac{1}{2}$ is relatively the same as for the electron. The splitting varies as Z^4 and is very small at low atomic numbers.

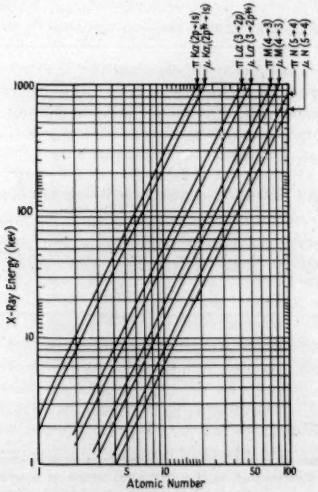


Figure 1. Calculated unperturbed energies of π and μ mesonic X-rays. (The fine structure terms are included only for the 1s and 2p levels and amount, at most, to 2·4% of the Bohr values within the range covered by the diagram.)

Energy values for some of the principal transitions, calculated from equations (4) and (5), are plotted in figure 1 for π and μ mesons. The principal corrections which must be applied to the energy values (4) and (5) are described in the following sections.

2.2. *Nuclear Size Effect*

The modification to the $-1/r$ Coulomb potential inside the nucleus can be taken account of at low Z by first-order perturbation theory. The energy shift with respect to the value calculated from equations (4) and (5) is then given by

† The gyromagnetic ratio of the μ meson has recently been shown to have a value very close to 2 as expected for a Dirac particle (Garwin *et al.* 1957, Cassels *et al.* 1957).

the usual relation†

$$\delta E = -e\int |\psi\psi^*|\,\delta V\,d\nu \qquad \ldots\ldots(6)$$

where δV is the difference between the actual potential and the Coulomb potential, ψ is the unperturbed wave function and e is the charge on the meson.

Cooper and Henley (1953) have evaluated (6) for the 1s state assuming a uniform charge distribution within the nucleus. Their expression may be generalized to all s states giving relation (7)

$$\frac{\delta E}{E_Z} = \frac{4}{5}\frac{1}{n^3}\left(\frac{ZR}{r_B}\right)^2 \qquad \ldots\ldots(7)$$

where R is the nuclear radius, r_B is the mesonic Bohr radius and E_Z is the binding energy of the 1s level. Equation (7) is only applicable to the K level for Z less than about 10 where the shift is approximately 1% but can be used to much higher Z values for $n > 1$.

The corresponding expression, equation (8), for the energy shift of the 2p level has been given by Flügge (1953, unpublished)‡ as

$$\frac{\Delta E}{E_Z} = \left\{0{\cdot}0018\left(\frac{ZR}{r_B}\right)^4 - 0{\cdot}0010\left(\frac{ZR}{r_B}\right)^5 + \ldots\right\}. \qquad \ldots\ldots(8)$$

The energy shift in this case is very small ($\sim 10^{-6} E_Z$ at $Z = 10$) since the wave function of p states is zero at $r = 0$.

The shift of the 1s level rapidly becomes appreciable for $Z > 10$ and the first-order corrections are no longer good enough. It is then necessary to solve the wave equation exactly using the appropriate potentials inside and outside the nucleus. The calculations have to be done numerically and have been carried out for a number of elements by Fitch and Rainwater (1953). Table 1 gives their calculated (μ-meson) values for the $2p_{3/2}$ and 1s energy levels for a nuclear radius, $R = 1{\cdot}3 \times 10^{-13} A^{1/3}$, together with the values given by the Dirac equation neglecting the nuclear size. The enormous effect of the nuclear size on the energy of the 1s level at high Z is discussed further in § 4.3 where the determination of nuclear radii from the energy of the K transitions is described.

Table 1. Calculated Effect of Nuclear Size on the Binding Energies (MeV) of the 1s and $2p_{3/2}$ levels

($m_\mu = 210 m_e$) (Fitch and Rainwater 1953)

	$R = 0$		$R = 1{\cdot}3 \times 10^{-13}\ A^{1/3}$	
Z	1s	$2p_{3/2}$	1s	$2p_{3/2}$
22	1·392	0·346	1·282	0·346
29	2·430	0·60	2·120	0·60
51	7·707	1·874	5·22	1·81
82	21·328	4·914	10·11	4·63

† The sign convention used throughout is that δE negative corresponds to an increase in binding energy and therefore a greater transition energy to the particular level.

‡ Lectures on field theory and applications to meson problems, at Carnegie Institute of Technology.

2.3. *Vacuum Polarization*

Quantum electrodynamics predicts a shift of the energy levels, from the values given by the Dirac equation, which has been verified in the case of electrons by Lamb and Retherford (1950). They observed an energy difference between the $2p_{1/2}$ and $2s_{1/2}$ levels in hydrogen (on the Dirac theory these levels are degenerate). The total frequency shift observed, $1057{\cdot}77 \pm 0{\cdot}10$ Mc/s, is to be compared with the calculated value $1057{\cdot}19 \pm 0{\cdot}16$ Mc/s of which $-27{\cdot}13$ Mc/s arises from the vacuum polarization term. The agreement between the theoretical and calculated shifts to within 0·6 Mc/s is therefore evidence that a vacuum polarization effect of $-27{\cdot}13 \pm 0{\cdot}6$ Mc/s is present. The main contribution to the Lamb–Retherford

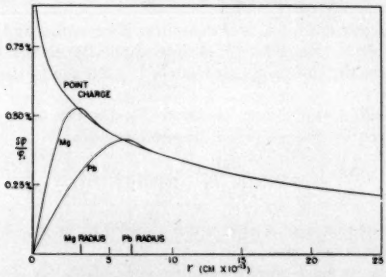


Figure 2. Change in electrostatic potential $\delta\phi$ close to an electric charge due to vacuum polarization (ϕ_C = Coulomb potential). (*After* Mickelwait and Corben 1954.)

shift for electrons arises from the radiation reaction term which, apart from its dependence on the dimensions of the wave function, varies as $1/m^2$ for other particles. Since the corresponding part of the vacuum polarization term is independent of mass, it gives the major contribution to the Lamb–Retherford shift in the case of mesons. The effect was first worked out by Uehling (1935) and the polarization of the vacuum is pictured as arising from the disturbance of the virtual electron† states in the immediate vicinity of an electric charge. The disturbance extends over a distance of the order of the Compton wavelength of the electron ($3{\cdot}9 \times 10^{-11}$ cm) away from an electric charge and within this region the effective charge increases slightly. Since the meson orbits are very close to the nucleus they will be influenced by this effect. The change of electrostatic potential close to an electric charge due to vacuum polarization has been plotted by Mickelwait and Corben (1954) and is reproduced in figure 2. These authors also calculated the corresponding level shifts in mesonic atoms using first-order perturbation theory (equation (6)). The calculated shifts for a point nucleus are plotted in figure 3 and amount at most to about 0·5%. The transition energies are *increased* by the vacuum polarization effect. The expected level shift due to the radiation reaction term for mesons is only a few per cent of that arising from vacuum polarization (Foldy and Eriksen 1954).

† Possible contributions to vacuum polarization from virtual states of mesons have been considered by Foldy (1954) and are negligible.

2.4. *Other Perturbations*

Apart from the nuclear force shift, which is specific to π mesons and is considered later, the remaining shifts to the mesonic energy levels are very small.

The part of the wave functions of the electrons which overlaps the meson's wave function will cause a reduction in the binding energy of the levels. The effect, which is least important for the tightly bound meson levels, has been estimated by Stearns and Stearns (1956). They find a fractional shift in energy for the $4f \rightarrow 3d$ transition in phosphorus of only 3×10^{-5}.

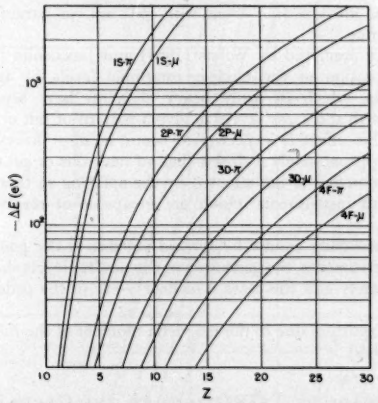


Figure 3. Calculated shifts of mesonic energy levels due to vacuum polarization (neglecting the nuclear size). (*After* Mickelwait and Corben 1954.)

An effect due to polarization of the nuclear charge by the meson when it is near the nucleus has been considered by Cooper and Henley (1953) and Lakin and Kohn (1954). In light elements the effect is expected to give a fractional increase in the transition energies of less than 10^{-4} but may approach approximately 10^{-3} for heavy elements.

2.5. *Fine Structure*

The fine structure separation for μ mesons corresponding to different j values for a given l, which is obtained from the Dirac formula, equation (5), is modified appreciably by nuclear size effects in high Z elements. In Pb for instance the $2p_{3/2}$–$2p_{1/2}$ separation is calculated to be 0·55 MeV from equation (5), but when the nuclear size is taken into account, Hill and Ford (1954) calculate that the separation should be between 0·176 MeV and 0·186 MeV.

There is no splitting of the levels of a given l for π mesons and so there is no fine structure of the K_α line ($2p \rightarrow 1s$). However the separations of the l substates for a given n are, according to equation (4), practically identical with the fine

structure separation for μ mesons. These separations which vary as Z^4 are very small in the light elements which can be studied in the case of π mesons. Moreover the effects of the nuclear interaction of π mesons will certainly modify this structure considerably.

There are additional effects which can also give splittings of the same order as the fine structure. Wheeler (1953) has predicted an effect due to the quadrupole moment of the nucleus. The quadrupole splitting arises from a variation of the nuclear size effect depending on the orientation of the **j** vector with respect to the nuclear angular momentum vector **I**. For heavy nuclei with $\mathbf{I} > \frac{1}{2}$ the $2p_{3/2}$ level may be split into as many as four levels with spacings comparable with the normal $2p_{3/2} - 2p_{1/2}$ difference.

A second effect predicted by Wilets (1954) and Jacobsohn (1954) arises from the Coulomb excitation of the nuclear rotational levels by the bound meson. The nuclear levels, which in many heavy elements have separations of about 50 kev, combine with the X-ray levels to give a pattern of ten or more lines in the K_α spectrum. This effect is expected to occur in most heavy elements except those containing approximately a closed shell of neutrons or protons. The centre of gravity of the structure is not affected by the splitting so that measurements of peak positions with instruments which are incapable of resolving the structure are not affected.

Although the structure cannot be studied yet due to the poor resolving power of the detecting instruments, the excitation of the nuclear levels should be detectable since in approximately half the cases a nuclear γ-ray of the order of 50 kev would be emitted.

The hyperfine splitting due to the magnetic moment of the nucleus is negligibly small (Wheeler 1949).

§ 3. Cascade of Transitions to the Ground State

The initial state into which the meson is captured by an atom will be influenced by the electron cloud. Once the meson is inside it however the cascade can be predicted with some certainty. The principal quantum number of the meson level having the same orbit radius as the K electron can be obtained from equation (2) and is given by

$$n = (\bar{\mu}/m_e)^{1/2}. \qquad \ldots\ldots(9)$$

Hence $n = 14$ for a μ meson and $n = 17$ for a π meson. Consequently the majority of the cascade can be treated without considering screening by the electrons. It is usually assumed that initially the various l states are occupied in proportion to their statistical weights $2l+1$ and that transitions having $\Delta n = 1$, occurring either by Auger electron emission or by radiation, are predominant. Detailed calculations have been performed by Wheeler (1949), Burbidge and de Borde (1953), de Borde (1954) and Demeur (1956). The main characteristics of the competition between radiation and Auger emission can be seen from figure 4 in which both transition rates calculated by Burbidge and de Borde are plotted as a function of Z for μ mesons. Figure 4 refers to transitions of the type $(n = n_1,\ l = n_1 - 1) \rightarrow (n = n_1 - 1,\ l = n_1 - 2)$, which are predominant, and the rates are plotted for various values of n. While the radiative transition rate varies

as Z^4 the Auger probability is nearly independent of Z but becomes more important for the higher n transitions. Except in the case of the K series, where radiation always predominates, Auger transitions predominate at low Z from a given level and are relatively more frequent from the higher levels.

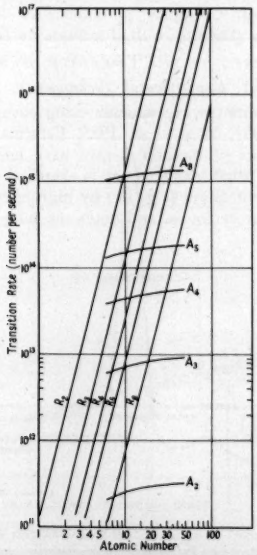


Figure 4. Calculated radiative (R_n) and Auger (A_n) transition rates for μ mesonic atoms (Burbidge and de Borde 1953). The radiative transition rate varies as Z^4, as plotted, only in the region where the nuclear size does not affect the transition energies appreciably (Wheeler 1949).

Hence a high yield of x-rays is expected in the K series for all Z and in the L series for $Z > 4$.

With π mesons the nuclear capture rate from the 2p level is comparable with the radiative rate. Messiah and Marshak (1952) have estimated that the rate of capture of π mesons from the 2p state is very approximately given by

$$w_{abs} \simeq 5 \times 10^{10} \times Z^6 \text{ sec}^{-1} \quad (Z > 1). \qquad \ldots\ldots(10)$$

The corresponding radiative transition probability is

$$w_r = 1{\cdot}75 \times 10^{11} Z^4 \text{ sec}^{-1}. \qquad \ldots\ldots(11)$$

Hence

$$w_{abs}/w_r \simeq 0{\cdot}3 Z^2 \qquad \ldots\ldots(12)$$

and nuclear capture is expected to predominate for $Z > 2$.

Hence the yield of K radiation will be very small in the case of π mesons except in the very light elements.

Experimental data on the Auger and radiative yields from π and μ mesonic atoms is available and is discussed in § 5. While the main features of these predictions about yields are confirmed, there is a marked disagreement at low atomic numbers.

§ 4. Measurements of the Energies of Mesonic X-rays

4.1. *Experimental Techniques*

Apart from a few qualitative experiments using cosmic-ray μ mesons (Chang 1949, 1954, Butement 1953, West *et al.* 1955, Fafarman and Shamos 1955) all the accurate measurements of mesonic x-rays have been made with artificially produced mesons. The usual arrangement is sketched in figure 5. π^- mesons produced in a light element target (e.g. Be) by high energy protons are deflected in the fringing field of the cyclotron and enter the long collimator in the shield.

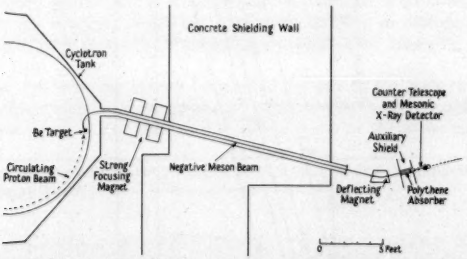


Figure 5. The meson beam from the Liverpool synchrocyclotron.

This defines the momentum of the particles and a deflecting magnet at the far end of the collimator prevents neutral radiation from entering the apparatus. Hence a meson beam of uniform momentum is obtained but it contains, in addition to π^- mesons, μ^- mesons (5–10%) from the decay of π mesons and electrons ($\sim$5%) from the decay of π^0 mesons near the target. For a π^- meson energy of 100 Mev, the energies of μ mesons and electrons are 116 Mev and 194 Mev respectively. The π and μ mesons can be completely separated because of the difference in their ranges.

Figure 6 shows a differential range curve in copper (Stearns and Stearns 1957 a), showing peaks from π and μ mesons. Approximately one third of the π^- mesons are stopped in a layer 2·5 g cm^{-2} in thickness at the optimum thickness of absorber. The targets cannot be made much thicker than 2 g cm^{-2} because of Compton scattering which degrades the radiation. Indeed it is fortunate that such thick targets can be used at all. This is due to the fact that the mesonic x-rays are so much more energetic than the ordinary x-rays from the same element that self-absorption due to photoelectric effect is negligible. A rather higher percentage

of mesons (up to 60%) can be stopped in the same target thickness if polythene is used as the slowing material instead of copper since the range straggling is smaller in polythene on account of the smaller degree of Coulomb scattering.

With present intensities of mesons the only practical methods of measuring the energies of the x-rays involve the direct use of scintillation counters or proportional counters. Crystal diffraction or magnetic deflection methods, using converters, are impractical on account both of intensity and the size of the meson beams (several square inches in cross section).

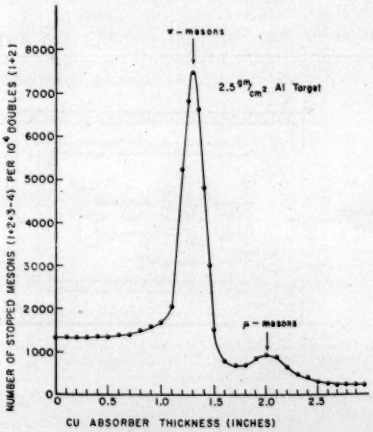


Figure 6. Differential range distribution of the meson beam from a synchrocyclotron. (*After* Stearns and Stearns 1957 a.)

The counter telescope arrangement used by Fitch and Rainwater (1953) to select pulses from mesonic x-rays has been adopted by all subsequent workers and is shown in figure 7. Scintillation counters 1, 2, and 3 select mesons stopped in the sample after they have been slowed down by the absorber. Counters 1 and 2 are connected in coincidence and counter 3 is in anti-coincidence with them. An event (12–3), in coincidence with a pulse from the detector, opens a gate a few microseconds wide which allow the pulse from the NaI scintillation counter to pass to an electronic pulse height analyser. The main disadvantage of this arrangement, which cannot however be avoided with present intensities, is that the detector is placed in the beam of particles and consequently counts the electrons and mesons which penetrate the sample. The resulting background rate in the detector limits the counting rate which can be accepted to a few counts per cyclotron pulse (usually about 100 μsec duration). At higher counting rates a variation of multiplier gain with counting rate is observed with scintillation counters and pile up of pulses occurs with proportional counters.

Most of the energy and intensity measurements have been made with scintillation counters which can detect radiations of the highest energies (6 MeV) from mesonic atoms. At energies below 100 kev the energy resolution of the scintillation counter is rather poor (the full width at half maximum is 25% at 60 kev for the best NaI crystals). In this energy region xenon-filled proportional counters (West and Bradley 1956) have a distinct advantage over scintillation counters in energy resolution (the full width at half maximum is 5% at 60 kev in the best conditions) but their efficiency falls off very rapidly with energy. In their usable energy range x-ray counting rates of 20–50 counts per minute can be obtained with proportional counters whereas the corresponding rate with scintillation counters is usually about 200 counts per minute. The proportional counter spectra are also subject to a smaller background presumably because of the much smaller stopping power of the counters for ionizing radiations.

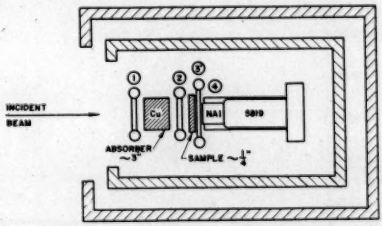


Figure 7. Counter telescope and NaI detector used by Fitch and Rainwater to measure mesonic x-rays. (*After* Fitch and Rainwater 1953.)

4.2. *Measurements of Low Energy X-rays and the Determination of the Meson Masses*

The object of the energy measurements to be described here was a determination of the π- and μ-meson masses. Transitions were chosen for which the corrections to the calculated energies, due to the Coulomb effects of finite nuclear size and, in the case of π mesons, the nuclear interaction were very small. The x-rays studied were all in the low energy region and the critical absorber technique was used. At an energy of 50 kev, the critical absorption edges of successive elements differ by 1·7 kev so that any individual line cannot be determined with much accuracy. However, by choosing the transitions studied to fit the energies of absorption edges available, it is possible to set narrow limits on the meson masses.

Koslov *et al.* (1954) studied μ meson transitions by this method. Their results are given in table 2. The mass values are calculated from the Dirac equation. A correction for finite nuclear size obtained from equation (7) was needed only in the case of the carbon $2p \rightarrow 1s$ transition and amounted to 0·6% of the transition energy.

The best independent value for the μ-meson mass is $206{\cdot}9 \pm 0{\cdot}3\ m_e$ for the μ^+ (Barkas *et al.* 1956), which conflicts with the limit $> 207{\cdot}55 m_e$ given by the measurement of the phosphorus line if vacuum polarization is neglected. Accordingly the measurements on the phosphorus line provide evidence for the existence

of a vacuum polarization shift in mesonic spectra. The limits on the μ-meson mass are then

$$206{\cdot}77m_e < m_\mu < 208{\cdot}95m_e. \qquad \ldots\ldots(13)$$

Table 2. Limits on the μ-meson Mass from Mesonic X-rays

	(1)	(2)	(3)	(4)
C	$2p \rightarrow 1s$	$<76{\cdot}123$	$<209{\cdot}99$	$<208{\cdot}95$
P	$3d \rightarrow 2p$	$>88{\cdot}015$†	$>207{\cdot}55$	$>206{\cdot}77$
Si	$4f \rightarrow 3d$	$>26{\cdot}713$	$>206{\cdot}82$	$>206{\cdot}47$

(1) Transition ; (2) energy determination (kev) ; (3) mass value (m_e) *without* vacuum polarization correction ; (4) mass value (m_e) *with* vacuum polarization correction.

† Koslov *et al.* used a value 88·065 kev for the absorption edge of Pb but Cohen, Crowe and DuMond (1956) recommend the value 88·015 kev as given here.

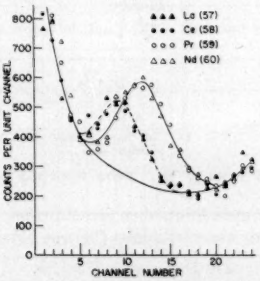


Figure 8. Pulse size distributions from the π mesonic M line of phosphorus with various critical absorbers. (*After* Stearns, Stearns, DeBenedetti and Leipuner 1954 a.)

Similar measurements with π mesons were carried out by Stearns, Stearns, DeBenedetti and Leipuner (1954 a), who restricted their measurements to $4f \rightarrow 3d$ transitions (i.e. the most probable $4 \rightarrow 3$ transition expected theoretically) for which the shift due to the nuclear interaction is negligible. Spectra from the phosphorus M line, with various critical absorbers interposed, are shown in figure 8. When the radiation is strongly absorbed, the peak is shifted to a lower energy ($\simeq 0{\cdot}75$ of the original value) corresponding to the fluorescent electronic K x-rays excited in the critical absorber.

The results of the measurements on π mesonic transitions are given in table 3.

Table 3. Limits on the π-meson Mass from Mesonic X-rays

	(1)	(2)	(3)	(4)
P	$4f \rightarrow 3d$	$>40{\cdot}44$	0·10	$>272{\cdot}2$
Al	$4f \rightarrow 3d$	$<30{\cdot}49$	0·065	$<273{\cdot}51$
K	$4f \rightarrow 3d$	$<65{\cdot}35$	0·190	$<273{\cdot}52$

(1) Transition ; (2) energy determination (kev) ; (3) vacuum polarization correction (kev) ; (4) mass of π meson (units of m_e).

In deriving the mass values, the Klein–Gordon equation was used and vacuum polarization corrections given in column (3) were included. Nuclear size effects are negligible. Hence the π mesonic x-ray measurements give

$$272{\cdot}2m_e < m_\pi < 273{\cdot}51m_e. \qquad \ldots\ldots(14)$$

Cohen, Crowe and DuMond (1956) have pointed out that even closer limits can be set on the π- and μ-meson masses by combining (13) and (14) making use of the result of Barkas *et al.* (1956) that

$$m_{\pi^+} - m_{\mu^+} = 66{\cdot}41 \pm 0{\cdot}1m_e. \qquad \ldots\ldots(15)$$

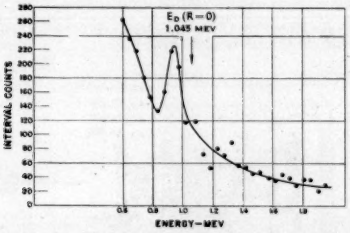


Figure 9. μ K x-rays of Ti_{22}. (*After* Fitch and Rainwater 1953.)

For instance, the *μ-meson* mass limits then give $273{\cdot}18m_e < m_\pi < 275{\cdot}36m_e$ and combining this with the directly observed limits (14) one obtains

$$273{\cdot}18m_e < m_\pi < 273{\cdot}51m_e \qquad \ldots\ldots(16)$$

with an uncertainty of $\pm 0{\cdot}1m_e$ on the mass limits. In a similar way one obtains, for the μ meson

$$206{\cdot}77m_e < m_\mu < 207{\cdot}1m_e. \qquad \ldots\ldots(17)$$

The limits (16) and (17) were used by Cohen, Crowe and DuMond in deducing their best values for the masses, which now stand at $m_\pi = 273{\cdot}27 \pm 0{\cdot}11m_e$ and $m_\mu = 206{\cdot}86 \pm 0{\cdot}11m_e$ (Crowe 1957).

4.3. *Measurements of High Energy μ-mesonic X-rays and the Determination of Nuclear Radii*

Fitch and Rainwater (1953) studied the energies of μ mesonic K x-rays in elements of intermediate and high atomic number with the object of measuring nuclear radii. Typical spectra, obtained with an NaI scintillation counter, from Ti_{22}, Cu_{29} and Pb_{82} are shown in figures 9, 10 and 11. In the lead spectrum three peaks are expected, separated by 0·51 MeV, corresponding to the re-absorption of 0, 1 or 2 quanta of annihilation radiation from the positrons produced by pair production from the incident mesonic x-rays. Fitch and Rainwater point out that their lead spectrum is not consistent with this interpretation and suggest that the peak at 5·3 MeV is the pair plus one annihilation quantum peak of a radiation of 5·8 MeV. The peak at 5·0 MeV is then interpreted as the pair peak

from a 6·0 MeV radiation. This would correspond to the $2p_{3/2} \rightarrow 1s$ transition whereas the 5·8 MeV radiation is interpreted as due to the $2p_{1/2} \rightarrow 1s$ transition. If this interpretation is correct it is difficult to see why a prominent peak is

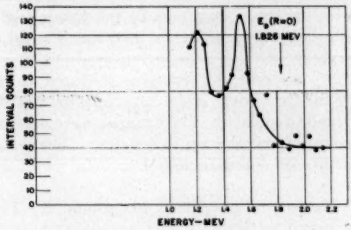


Figure 10. μ K x-rays of Cu_{29}. (*After* Fitch and Rainwater 1953.)

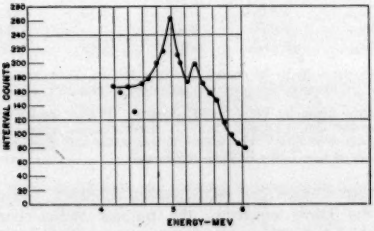


Figure 11. μ K x-rays of Pb_{82}. (*After* Fitch and Rainwater 1953.)

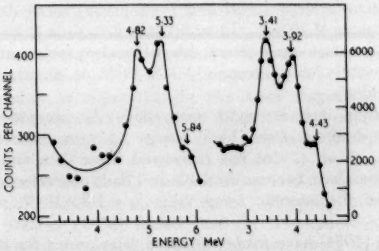


Figure 12. Pulse height distributions from μ K x-rays of Pb (left) and from a nuclear gamma ray (right) of energy 4·43 MeV. (*After* Butement 1955.)

not present at 5·5 MeV from the 6·0 MeV radiation, since it should be twice as intense as the postulated $2p_{1/2} \rightarrow 1s$ peak at 5·3 MeV. The lead spectrum has been re-measured by Butement (1955) using a thicker NaI crystal than Fitch and Rainwater's. Butement's spectrum is shown in figure 12. The pair plus one

annihilation quantum peak at 5·33 MeV is more intense relative to the pair peak than in figure 11, but the spectrum shows only the three peaks and one must conclude that there is no definite evidence that the $2p_{3/2} \rightarrow 1s$ and $2p_{1/2} \rightarrow 1s$ transitions have been resolved.

The point is fortunately not important in the bearing of the measurements on nuclear radii since the energy difference involved is small compared with the very large shifts observed.

The experimentally observed K_α energies (reduced to $2p_{3/2} \rightarrow 1s$ transition energies by assuming the theoretical $2p_{3/2} - 2p_{1/2}$ separation of the unresolved lines in the K_α peak) are listed in table 4. The energies were determined from calibrations with nuclear γ-rays of energies up to 4·4 MeV. The accuracy of the energy measurements is stated to be approximately 1%.

Table 4. Measured Energies of μ mesonic K_α Transitions

(1)	(2)	(3)	(4)	(5)
Ti_{22}	1·046	0·955		1·17
Cu_{29}	1·828	1·55	1·55	1·21
Sb_{51}	5·833	3·50		1·22
Pb_{82}	16·414	6·02†	5·89	1·17†

(1) Element ; (2) calculated $2p_{3/2} \rightarrow 1s$ energy (MeV) for $R = 0$; (3) observed $2p_{3/2} \rightarrow 1s$ energy (MeV) (Fitch and Rainwater) ; (4) observed $2p_{3/2} \rightarrow 1s$ energy (MeV) (Butement) ; (5) r_0 (units of 10^{-13} cm).

† These are the values given by Fitch and Rainwater. If their peak at 5·0 MeV is taken as a composite peak due to the $2p_{3/2} \rightarrow 1s$ and $2p_{1/2} \rightarrow 1s$ transitions, as seems probable, the $2p_{3/2} \rightarrow 1s$ transition energy becomes 6·08 MeV. The mean of this value and Butement's is then 5·98 MeV, corresponding to a value of $r_0 = 1{\cdot}18 \times 10^{-13}$ ($\mu = 210\ m_e$).

Other elements were studied but with poorer accuracy. Fitch and Rainwater evaluated, from the Dirac equation, the nuclear radius constant $r_0 = RA^{-1/3}$ required to give the observed energies assuming a uniform charge distribution in the nucleus. The values obtained using a meson mass $m_\mu = 210 m_e$ are given in the last column of table 4. Hill and Ford (1954) recalculated the value for Pb using a meson mass of $207 m_e$. In addition to a vacuum polarization correction, they also included other small terms due to nuclear polarization and radiative effects but their value for r_0, namely $1{\cdot}18 \times 10^{-13}$, is practically identical with Fitch and Rainwater's.

At the time these measurements were made the accepted value of r_0 was $1{\cdot}4 \times 10^{-13}$ cm as deduced from high energy electron scattering and nuclear physical methods. Scott (1956) has reassessed these data and the more recent determinations which have become available and finds no evidence for a discrepancy between them and the mesonic x-ray value $r_0 = 1{\cdot}2 \times 10^{-13}$, which is now the accepted value.

Hill and Ford (1954) have made extensive calculations for lead on the effects of the shape of the nuclear charge distribution on mesonic x-ray energies. They find that the energy shift of the $2p \rightarrow 1s$ transition depends mainly on the average value of $r^{0.5}$ and that a wide variety of shapes of distributions can give the observed energy. The $3d \rightarrow 2p$ and $2s \rightarrow 2p$ transition energies are however more sensitive to the shape of the charge distribution and it is possible that an extensive study of μ mesonic x-ray spectra would give useful information about the distribution.

4.4. *π-mesonic X-rays and the Nuclear Force Shift*

Unlike many of the phenomena associated with mesonic atoms, the nuclear force shift was observed experimentally (Stearns, Stearns, DeBenedetti and Leipuner 1954 b, Stearns, Stearns, Leipuner and DeBenedetti 1954) for the K_α transitions before it had been predicted theoretically. Shortly afterwards calculations of the shift of the 1s level were made by Deser *et al.* (1954) and Brueckner (1955). The former theory connects the shift with the elastic interaction between π mesons and single nucleons described by the s wave scattering lengths a_1 and a_3. Brueckner deduced an additional shift due to the inelastic interaction of π mesons with two nucleons which leads to the well-known absorption of π^- mesons.

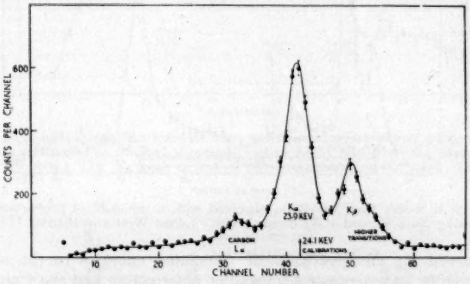


Figure 13. Pulse height distribution of π mesonic K x-rays from lithium measured with a proportional counter. (*After* West and Bradley 1957.)

The additional shift predicted by Brueckner is approximately 0·7 times that according to Deser *et al.* A later extension of Brueckner's theory by Karplus and Halpern (1957) gives a smaller total additional shift, 0·25 that according to Deser *et al.* Finally, Wolfenstein (1957) has included the elastic p wave interaction which can contribute to the scattering in the 1s state due to the motion of the nucleons in the nucleus, and gives a shift of opposite sign to that from the inelastic interaction. This shift is approximately the same magnitude as the inelastic shift deduced by Karplus and Halpern. There is consequently considerable interest in accurate measurements of the energy shifts for comparison with these theories.

Initial measurements of the K_α energies using critical absorbers were rapidly replaced by more accurate direct measurements. However, in two cases, Be and ^{10}B, Stearns and Stearns (1956) found that the absorption edge of a critical absorber fell so close to the position of the line that, owing to the natural width of the line (§ 5.2), partial absorption occurred. In these cases the critical absorber measurements give a useful check on the pulse height measurements.

Both scintillation counters (Camac, McGuire, Platt and Schulte 1955, Stearns and Stearns 1956) and proportional counters (West and Bradley 1957) have been used to measure the energies of the K_α transitions. Proportional counter spectra for Li, Be and ^{10}B K x-rays are shown in figures 13 to 15 and show $K_\alpha(2p \rightarrow 1s)$

and K_β(3p→1s, 4p→1s etc.) radiations clearly resolved. Above boron the proportional counter is too inefficient for accurate measurements. However the scintillation counter can be used for measurements up to F although for the lower atomic number elements (below B) the K_β radiation is not properly resolved. A scintillation counter spectrum of K x-rays from N (Stearns and Stearns) is shown in figure 16.

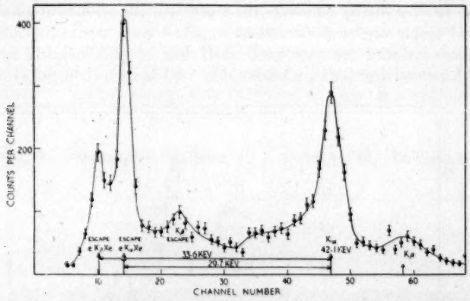


Figure 14. π K x-rays from beryllium observed with a xenon-filled proportional counter showing Xe e K_α and e K_β escape peaks. (*After* West and Bradley 1957.)

The observed K_α energies and the calculated energies in the absence of a nuclear force shift, i.e. corrected for vacuum polarization and the Coulomb effects of nuclear size only, are given in table 5.

Table 5. Measured Energies of π-mesonic K_α Transitions

(1)	(2)	(3)	(4)	(5)	(6)	(7)
^{6}Li	24·57			24·31 ± 0·1	24·31 ± 0·1	0·9 ± 0·4
^{7}Li	24·65		23·73† ± 0·12	23·85 ± 0·1	23·80 ± 0·08	3·3 ± 0·3
^{9}Be	44·03	43 ± 3·5	42·09 ± 0·10	42·15 ± 0·15	42·11 ± 0·07	4·2 ± 0·16
^{10}B	68·89		65·2 ± 0·2	65·3 ± 0·4	65·22 ± 0·18	5·1 ± 0·26
^{11}B	68·96		63·5 ± 0·2	64·5 ± 0·3	63·80 ± 0·4	7·4 ± 0·6
C	99·27	92 ± 7·5	92·6 ± 0·4		92·6 ± 0·4	6·5 ± 0·4
N	135·04	132 ± 2·5	126·0 ± 0·3		126·1 ± 0·3	6·6 ± 0·2
O	176·06	163·8 ± 2·7	155·2 ± 0·8		155·9 ± 0·8	11·4 ± 0·5
F	222·01		196·3 ± 1·0		196·3 ± 1·0	11·4 ± 0·5

(1) Element or isotope; (2) calculated energy E_0 (kev), $m_\pi = 273{\cdot}3m_e$; (3)–(5) measured energy E (kev): (3) Camac, McGuire *et al.* (1955), scintillation counter, (4) Stearns and Stearns (1956), scintillation counter, (5) West and Bradley (1957) proportional counter; (6) weighted mean energy E (kev); (7) % shift $(E_0-E)/E_0$.

† Corrected for ^{6}Li content of natural lithium.

The critical absorption measurements of Stearns and Stearns show a partial absorption of the Be line in Pr (absorption edge 41·98 kev) and of the ^{10}B line in Hf (absorption edge 65·32 kev) confirming the energy values measured directly. There are significant differences between the values measured by different observers in the cases of ^{11}B, N and O but as there is no reason for preferring a particular

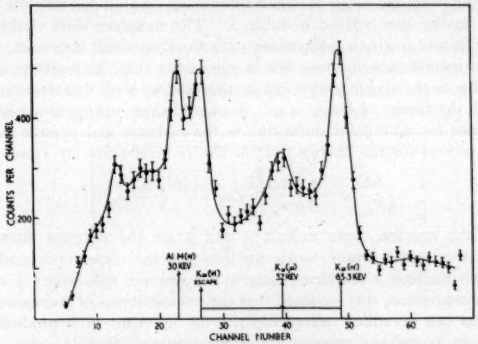


Figure 15. π K x-rays from ^{10}B observed with a five-atmosphere xenon proportional counter. Prominent peaks due to residual μ mesons and to Al M rays from the thin box containing the ^{10}B are seen. The K_β peak is lost in the background in this case. (*After* West and Bradley 1957.)

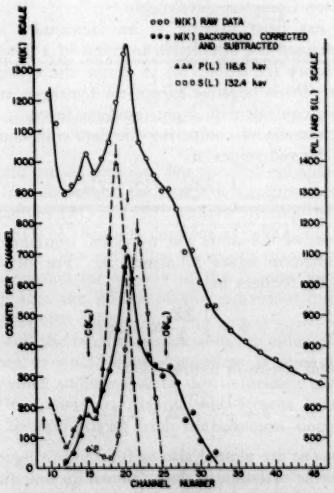


Figure 16. π K x-rays from nitrogen observed with an NaI scintillation counter. Peaks due to π K x-rays of C and O (from a small contamination of water in the hydrazine used as a target) are also present. (*After* Stearns and Stearns 1956.)

measurement a weighted mean has been taken in all cases in deriving the percentage shifts given in the last column of table 5. The measurements clearly establish an energy shift and justify a comparison with the theoretical estimates.

The theories mentioned above are in agreement that the major component of the shift is due to the elastic s wave interaction and we shall therefore first compare the data with the theory of Deser *et al.* It should then be apparent whether there is any evidence for additional shifts due to the inelastic and p wave interactions.

The expression for the energy shift of the 1s level given by Deser *et al.* is

$$\frac{\delta E}{E_{K_\alpha}} = -\frac{4}{3}\frac{\mu}{\bar{\mu}}\frac{4Z}{r_B}\left\{\frac{2}{3}Za_1+\frac{(3N+Z)}{3}a_3\right\} \qquad \ldots\ldots(18)$$

where r_B is the mesonic Bohr radius, μ and $\bar{\mu}$ are the reduced masses for the π meson–nucleus and π meson–single nucleon systems respectively and a_1 and a_3 are the meson–nucleon s scattering lengths for isotopic spin states $\frac{1}{2}$ and $\frac{3}{2}$. In deriving this expression, it is assumed that the contributions of individual nucleons in the nucleus can be added independently, the neutrons each producing a shift proportional to a_3 and the protons a shift proportional to $\frac{1}{3}(2a_1+a_3)$. The best values of a_1 and a_3 have been derived by Orear (1956) from a least squares analysis of all the scattering and photo production data.

He obtains $$a_1 = 0{\cdot}167 \pm 0{\cdot}012\lambda_\pi, \quad a_3 = -0{\cdot}105 \pm 0{\cdot}010\lambda_\pi \qquad \ldots\ldots(19)$$

where λ_π is the π meson Compton wavelength.

Expression (18) has been derived in an elementary way by Bethe and de Hoffman who expressed the interaction in terms of a potential and used first-order perturbation theory (equation (6)) to derive the energy shift. The term in braces in equation (18) is negative except for hydrogen and corresponds to a raising of the 1s level, equivalent to a net repulsive force.

The most instructive way of comparing the data with equation (18) is due to Karplus (1956) who derived values of

$$\delta = \frac{1}{\lambda_\pi}\left\{\frac{2Za_1+(3N+Z)a_3}{3}\right\} \qquad \ldots\ldots(20)$$

from the measurements of the shifts for individual isotopes using equation (18). These values are plotted in figure 17 against Z. For $N = Z$ (open circles in figure 17) equation (20) reduces to

$$\delta = \frac{1}{\lambda_\pi}\frac{2Z}{3}(a_1+2a_3) \qquad \ldots\ldots(21)$$

and for $N = Z+1$ (solid circles in figure 17)

$$\delta = \frac{1}{\lambda_\pi}\left\{\frac{2Z}{3}(a_1+2a_3)+a_3\right\}. \qquad \ldots\ldots(22)$$

These two expressions are plotted also as full lines in figure 17. The broken lines indicate the possible variation of slope (shown for one direction only) due to the errors in a_1 and a_3 which are enhanced in the combination $\frac{2}{3}(a_1+2a_3)$ which gives the slope of the lines. Very recent measurements of a_1 and a_3 by Barnes, Rose, Ring and Jiacomelli (1957, private communication) ($a_1 = 0{\cdot}169 \pm 0{\cdot}022\lambda_\pi$,

$a_3 = -0{\cdot}106 \pm 0{\cdot}003\lambda_\pi$) have been combined with Orear's values in plotting the lines in figure 17.

It is evident that there is very little agreement between the experimental data and equations (21) and (22). At very low Z there is perhaps a tendency to follow the straight lines but large and significant deviations occur between one element and the next. The situation appears to be even less satisfactory at higher Z.

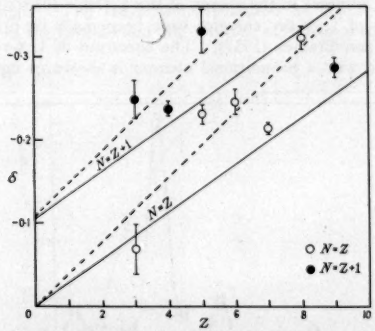


Figure 17. Comparison of the experimental shifts of the π mesonic K x-ray line with the theory of Deser *et al.* The experimental points are derived from the relation.

$$\delta = -\frac{3}{16}\frac{\bar{\mu}}{\mu}\frac{r_B}{Z}\frac{\delta E}{E_{K_\alpha}}\frac{1}{\lambda_\pi}$$

and the two parallel lines represent the variation of δ according to the theory of Deser *et al.* for nuclei with $N = Z$ and $N = Z+1$ respectively. (*After* Karplus 1956.)

Karplus (1957, private communication) has pointed out that the additional shifts due to the inelastic interaction and the elastic p wave interaction will merely alter the constants in equations (21) and (22) without altering the form of the relation and that linear relations with the same slope for nuclei having $N = Z$ and $N = Z+1$ would still be expected. The individual variations from one element to the next are therefore not accounted for by any of the current theories, and it appears that the assumption that the effects of the individual nucleons can be added independently is not strictly valid.

In view of this, and the large uncertainty in the value of $a_1 + 2a_3$ which gives the slope of the lines, it would appear pointless at present to enquire whether there is evidence for small additional shifts due to inelastic and p wave interactions.

Wolfenstein (1956, quoted by Stearns and Stearns 1956) has estimated the energy shift of the 2p level arising from the π meson–single nucleon interaction and obtains a shift

$$\frac{\delta E}{E_{2p}} = -\frac{\mu}{\bar{\mu}}(Z\alpha)^3\frac{\delta_{33}}{\eta^3}\left(\frac{Z}{3}+N\right) \qquad \ldots\ldots(23)$$

where E_{2p} is the binding energy of the 2p level and δ_{33} is the p wave phase shift corresponding to $j = \frac{3}{2}$, $T = \frac{3}{2}$, which is the largest of the p wave phase shifts,

and η is the centre-of-mass momentum in units of $m_\pi c$. The contribution of the other phase shifts may not be negligible and this expression only gives an idea of the magnitude of the expected shift. Using $\delta_{33} = 0{\cdot}23\eta^3$ (Orear 1956) the expression reduces to

$$\frac{\delta E}{E_{L_\alpha}} = -2{\cdot}5 \times 10^{-5} Z^4 \% \qquad \ldots\ldots(24)$$

which gives a 4% increase in the energy of the $L_\alpha(3d \rightarrow 2p)$ transition for $Z = 20$.

Measurements of L_α x-ray energies have been made by Stearns and Stearns (1956) and West and Bradley (1957). The spectrum of L x-rays from nitrogen ($C_2N_4H_4$) obtained with a proportional counter is shown in figure 18 and shows

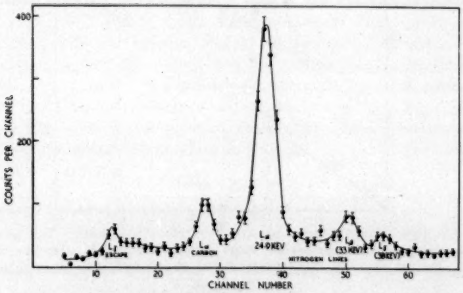


Figure 18. π L x-rays from nitrogen ($C_2N_4H_4$) observed with a proportional counter. (*After* West and Bradley 1957.)

resolved groups due to the $L_\beta(4 \rightarrow 2)$ and $L_\gamma(5 \rightarrow 2)$ transitions, as well as a smaller peak due to C L_α x-rays. At $Z = 20$ Stearns and Stearns observed a shift of $1 \pm 1\%$ and so it would appear that equation (24) overestimates the shift considerably. Below $Z = 10$ the shifts are less than 0·2% according to equation (24) and measurements of L_α energies in elements from C to Mg were used by West and Bradley to obtain a value for the π-meson mass neglecting nuclear force shifts altogether. The value obtained, $m_\pi = 272{\cdot}5 \pm 0{\cdot}8 m_e$, is very close to the accepted value and provides additional evidence that the L_α transition energies are not altered by more than a few tenths of one per cent in the light elements. There is a possibility that appreciable shifts of the 2p level will occur in the heavier elements, but no measurements have yet been made above $Z = 20$.

§ 5. Experimental Determination of the Yields of K and L series X-rays

5.1. *X-ray Yields from μ mesonic Atoms*

According to the calculated values of the Auger and radiative transition probabilities plotted in figure 4, it was expected that the K x-ray yield would be nearly 100% in all elements for μ mesonic x-rays and that the L x-ray yield would only differ appreciably from 100% below $Z = 10$. Qualitative support for these

conclusions was provided by the measurements of Auger electron yields in photographic emulsions. The most recent measurements of Fry (1953) gave the Auger yields as 4·7% for the light elements in the emulsion (C, N and O) and 44% for the Ag and Br. The corresponding calculated values (for electrons in the energy range 10–100 kev) are 1·6% and 100% (Burbidge and de Borde 1953).

Measurements of the yields of K and L series x-rays by Stearns and Stearns (1957 a) however showed a marked disagreement with these predictions in the case of the light elements. The yields were measured by counting the number of mesons stopping in the sample and the number of x-rays recorded in the peaks

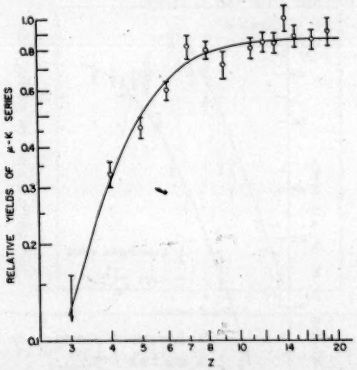


Figure 19. Total yield of μ K x-rays plotted against Z. The curve is drawn using $\omega_{coll} + \omega_{Auger} = 5{\cdot}9 \times 10^{13}$ sec^{-1}, $\omega_r = 1{\cdot}3 \times 10^{11} Z^4$ sec^{-1}. The errors refer only to the relative yields, the vertical scale being assigned from the mean value of the absolute yield for O ($0{\cdot}8 \pm 0{\cdot}15$). (*After* Stearns and Stearns 1957 a.)

of the scintillation counter spectra. Careful determinations of the overall efficiency of the apparatus as a function of energy using radioactive sources then enabled the number of x-rays per meson stopped to be estimated. In figures 19 and 20 the experimental total yields for the K and L series x-rays are plotted as a function of Z. The reduced K x-ray yield at low Z and the fact that the L x-ray yield has not reached its maximum value at $Z = 10$ are both in disagreement with the calculations of the Auger competition.

The yield curves can be fitted by expressions of the type

$$Y = C_1 \frac{aZ^4}{C_2 + aZ^4} \qquad \ldots\ldots(25)$$

where C_1, C_2 and a are constants.

This is of the form expected for competition between a radiative transition probability $\omega_r = aZ^4$ and another process which is independent of Z. The radiative transition probabilities are readily calculated. In the case of the K series it is predominantly the $2p \rightarrow 1s$ transition probability which is involved and a is then $1{\cdot}3 \times 10^{11}$ sec^{-1}.

For the L series the $3d \rightarrow 2p$ transition is mainly involved and a is calculated to be $1{\cdot}3 \times 10^{10}$ sec^{-1}. The constants C_1 and C_2 are then chosen to fit the experimental data. The curves drawn in figures 19 and 20 employ the following values: for the K series

$$C_1 = 0{\cdot}86, \quad C_2 = 5{\cdot}9 \times 10^{13} \text{ sec}^{-1} \qquad \ldots\ldots(26)$$

and for the L series

$$C_1 = 0{\cdot}97, \quad C_2 = 2{\cdot}7 \times 10^{14} \text{ sec}^{-1} \qquad \ldots\ldots(27)$$

whereas for an Auger process one would expect $C_2 \simeq 2 \times 10^{11}$ sec^{-1} for the K series and $C_2 \simeq 8 \times 10^{12}$ sec^{-1} for the L series.

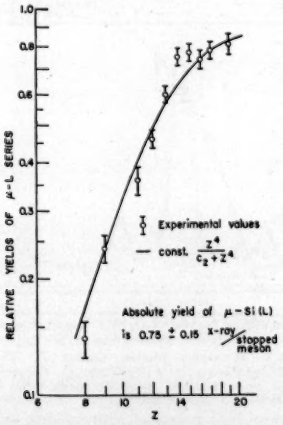


Figure 20. Total yield of μ L series x-rays plotted against Z. The curve is drawn using $\omega_{coll} + \omega_{Auger} = 2{\cdot}7 \times 10^{14}$ sec^{-1}, $\omega_r = 1{\cdot}3 \times 10^{10} Z^4$ sec^{-1}. The errors refer to the relative yields only. (*After* Stearns and Stearns 1957 a.)

Day and Morrison (1957) have suggested that collision de-excitation could account for the observed magnitude of C_2. As a result of the emission of Auger electrons, the excited atom acquires a recoil motion. During collisions with stationary atoms the mesonic atom could de-excite by ejecting electrons from the stationary atom. An approximate calculation gives $C_2 \simeq 10^{13}$ sec^{-1} for the 2p state in Li, varying only slowly with Z, which is much closer to the observed value than the free atom Auger coefficient. The authors point out that in an appreciable fraction of cases the meson should transfer to the stationary atom, a process which is in fact observed in liquid hydrogen (§ 8). A possible objection to this explanation is that it may require an Auger electron emission probability much higher than is observed in the light elements. Day and Morrison suggest that the hypothesis could be tested by measuring the yield of x-rays in a gaseous target as a function of pressure.

5.2. *X-ray Yields from π mesonic Atoms*

The K and L series x-ray yields from π mesonic atoms have been measured by Camac, McGuire *et al.* (1955), Camac, Halbert and Platt (1955), Stearns and Stearns (1957 b) and Stearns *et al.* (1957). The yields are plotted in figures 21 and 22 where only the data of Stearns *et al.* have been quoted in the L series since they are considerably the more accurate. In addition to a falling off in yield at low atomic numbers, similar to that observed with μ mesons, both the K and L series yields show ultimately a reduction of yield with increasing Z,

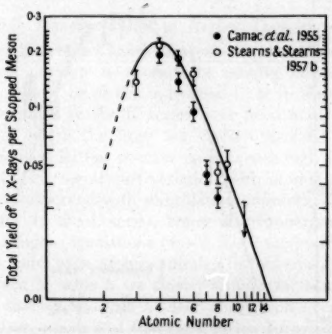


Figure 21. Yield of π mesonic K x-rays plotted against Z. The curve is drawn with $\omega_{\text{coll}}+\omega_{\text{Auger}} = 4\times10^{13}$ sec^{-1}, $\omega_{\text{r}} = 1{\cdot}75\times10^{11}Z^4$ sec^{-1}, $\omega_{\text{abs}} = 3{\cdot}5\times10^{9}Z^7$ sec^{-1}.

which can be accounted for by direct nuclear capture mainly from the $n=2$ and $n=3$ states respectively. The yield can be written

$$Y_{\text{r}} = P_{n,l}\frac{\omega_{\text{r}}}{\omega_{\text{coll}}+\omega_{\text{Auger}}+\omega_{\text{abs}}+\omega_{\text{r}}} \qquad \ldots\ldots(28)$$

where ω_{coll}, ω_{Auger}, ω_{abs} and ω_{r} are the collision, direct Auger, nuclear absorption and radiation probabilities from the state (n, l) and $P_{n,l}$ is the probability that the meson passes through the state (n, l). ω_{coll} and ω_{Auger} are slowly varying functions of Z. It is shown in the appendix (equation (A8)) that the probability of nuclear absorption from any state of orbital angular momentum l varies as $Z^{2l/3+4}$ for low atomic numbers ($Z<10$) whereas ω_{r} varies as Z^4.

Let us consider first the total L series yield (figure 22). The radiative transition probability for the $3d\rightarrow2p$ transition, which probably accounts for nearly 80% of the L yield, is calculated as

$$\omega_{\text{r, 3d}} = 1{\cdot}77\times10^{10}Z^4 \text{ sec}^{-1}. \qquad \ldots\ldots(29)$$

$\omega_{\text{coll}}+\omega_{\text{Auger}}$ can then be deduced from the variation of yield at low atomic numbers where ω_{abs} is negligible. The fit in figure 22 is obtained with

$$(\omega_{\text{coll}}+\omega_{\text{Auger}})_{\pi\text{L}} = 1{\cdot}8\times10^{14} \text{ sec}^{-1} \qquad \ldots\ldots(30)$$

which is similar to the value ($2{\cdot}7 \times 10^{14}$ sec^{-1}) obtained for the μ L series. The high Z region in figure 22 can then be fitted by a term

$$\omega_{abs} = 7{\cdot}1 \times 10^{4} Z^{8} \text{ sec}^{-1} \qquad \ldots\ldots(31)$$

although there are appreciable departures from the fitted curve in the region of $Z = 28$ which is a magic number nucleus. Such effects have been discussed by Messiah and Marshak (1952).

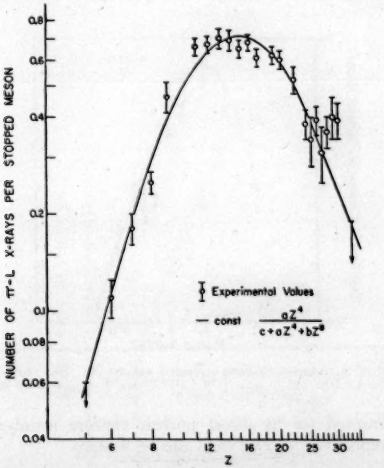


Figure 22. Total yield of π L x-rays plotted against Z. The curve is drawn using $\omega_\pi = 1{\cdot}77 \times 10^{10} Z^4$ sec^{-1}, $\omega_{coll} + \omega_{Auger} = 1{\cdot}8 \times 10^{14}$ sec^{-1}, $\omega_{abs} = 7{\cdot}1 \times 10^4 Z^8$ sec^{-1}. (*After* Stearns, Stearns and Leipuner 1957.)

The main components of the L-ray spectrum consist of d$\rightarrow$p transitions for which, according to the relation quoted above, the absorption rate should vary as $Z^{\sim 3}$ at low atomic numbers. The eighth power of Z occurring in equation (31), which is derived from data in the intermediate Z region, is therefore not inconsistent with this estimate. Equation (31) may, to a first approximation, be used to derive the nuclear absorption rate of π mesons from the 3d state.

The K x-ray yields (figure 21) can similarly be fitted by relation (28). $\omega_{coll} + \omega_{Auger}$ cannot be deduced independently of ω_{abs} as with the L series since the absorption term is important even at very low atomic numbers. Accordingly Stearns and Stearns (1957 b) chose

$$(\omega_{coll} + \omega_{Auger})_{\pi K} = 4 \times 10^{13} \text{ sec}^{-1} \qquad \ldots\ldots(32)$$

which gives the same ratio for $\omega_{coll} + \omega_{Auger}$ in the K series for π and μ mesonic atoms as obtains in the L series.

Combined with the calculated value

$$(\omega_r)_{2p} = 1{\cdot}75 \times 10^{11} Z^4 \text{ sec}^{-1} \quad \ldots\ldots(33)$$

for the $2p \rightarrow 1s$ radiative transition rate, this enables ω_{abs} to be deduced. The fit in figure 21 requires

$$\omega_{abs} = 3{\cdot}5 \times 10^9 Z^7 \text{ sec}^{-1} \quad \ldots\ldots(34)$$

whereas one would expect $\omega_{abs} \propto Z^{6\cdot7}$ for the 2p state (appendix, eqn (A5)). There is some evidence of an anomaly in the yield near $Z = 8$ which is a magic number nucleus.

5.3. *Relative Yields of Higher Transitions*

In the absence of a detailed knowledge of the cascade of transitions to the ground state, it is not possible to predict the relative intensities of $K_\alpha(2p \rightarrow 1s)$ and $K_\beta(3p \rightarrow 1s)$ etc. x-rays, or of the individual lines in the L series. Measurements of relative intensities in the K series have been made for Li and Be with proportional counters where the lines are clearly resolved and estimates have also been made from scintillation counter spectra although in this case the lines are not well resolved. No significant variation with atomic number of the yield of higher transitions is observed with scintillation counters, and mean values only are quoted in table 6. In the L series, better discrimination is possible between the $L_\alpha(3 \rightarrow 2)$ and the higher transitions ($4 \rightarrow 2$, $5 \rightarrow 2$ etc.) and again no significant variation of relative yield with atomic number is observed. The yields of the higher transitions given in table 6 are closely similar for π and μ mesonic atoms in spite of the fact that appreciable nuclear absorption occurs only in the case of π mesons. This fact, which is at first sight rather surprising, may be explained qualitatively as follows. Consider first the region of Z where nuclear absorption competes strongly with radiation from the initial state and Auger and collision de-excitation are negligible i.e., $Z > 5$ in the K series and $Z > 25$ in the L series.

Table 6. Relative Yields of the Higher Transitions in the K and L Series for π and μ mesonic Atoms

Elements	Authors	K_β + higher transitions / total K series		L_β + higher transitions / total L series	
		π mesons	μ mesons	π mesons	μ mesons
Li	WB 1957	$0{\cdot}32 \pm 0{\cdot}03$			
Be	WB 1957	$0{\cdot}21 \pm 0{\cdot}02$			
$Z = 5$ to $Z = 11$	SS 1957 b	0·23			
$Z = 4$ to $Z = 19$	SS 1957 a		0·23		
$Z = 6$ and 7	WB 1957			0·19	
$Z = 6$ to $Z = 16$	CHP 1955			0·25	
$Z = 7$ to $Z = 24$	SSL 1957			0·23	
$Z = 26$ to $Z = 30$	SS 1957 b			0·22	
$Z = 8$ to $Z = 19$	SS 1957 a				0·29

WB = West and Bradley ; SS = Stearns and Stearns ; CHP = Camac, Halbert and Platt ; SSL = Stearns, Stearns and Leipeuner.

In the K series, nuclear absorption from the 3p state (from which the K_β line originates) would be expected to be about one third as rapid (appendix, eqn (A7)) as

from the 2p state (from which the K_α line originates). The radiative transition rate from the 3p level (predominantly involving transitions to the 1s level) is also about one quarter of the radiative transition rate from the 2p state (Bethe and Salpeter 1957, p. 352). Consequently absorption competes with radiation to about the same extent from the 3p level as from the 2p level and so the relative intensities of the K_α and K_β lines will remain practically unaffected by nuclear absorption.

In the region of low atomic number where collision de-excitation is important, the effect on the higher transitions has not been worked out. Experimentally however there is some evidence of an increase in the ratio of higher transitions observed with lithium.

§6. Nuclear Capture from the 1s Level

6.1. μ *Mesons: Lifetime Measurements*

As mentioned in the introduction, the variation of the lifetime for decay of negative μ mesons with the atomic number of the stopping material constituted the first evidence for mesonic atom formation. Extensive measurements of the variation of lifetime with atomic number have been made since that time and the data constitute the most accurate information on the variation with Z of the nuclear capture rate from a mesonic Bohr level. It is not proposed however to discuss these measurements in detail and for further information the reader is referred to the review article by Sard and Crouch (1954).

The capture probability is given by

$$\omega_{\text{abs}} = \text{const.} \sum_{\text{all protons}} \psi(\text{at each proton})^2 \qquad \ldots\ldots(35)$$

where ψ is the mesonic wave function for the 1s level. For Z less than about 10 the wave function at the origin may be used for all the protons and since for hydrogenic wave functions, $|\psi_0|^2 \propto Z^3$ (appendix, eqn (A2)), one obtains

$$\omega_{\text{abs}} \propto Z^4. \qquad \ldots\ldots(36)$$

At higher Z the modification to the wave function due to the finite nuclear size and the variation of ψ over the nuclear volume must be considered. Wheeler (1949) has calculated that in this case

$$\omega_{\text{abs}} = \frac{1}{\tau_0}\left(\frac{Z_{\text{eff}}}{Z_0}\right)^4 \qquad \ldots\ldots(37)$$

where τ_0 is the mean life of the unbound μ^- meson, Z_0 is determined from experiment and

$$Z_{\text{eff}} = Z\left\{1+\left(\frac{Z}{37{\cdot}2}\right)^{1{\cdot}54}\right\}^{-1/1{\cdot}54}. \qquad \ldots\ldots(38)$$

This relation, together with the experimental values of the capture rate, is plotted in figure 23. Absorption at low Z follows relation (37) up to $Z \simeq 30$ but for higher Z the observed capture rates fall below the curve. Primakoff (1955) has explained this disagreement in terms of the high neutron excess in heavy elements which limits the number of states in which the neutron, formed as a result of the absorption of the μ^- meson, can be accommodated. His calculated values for several individual nuclei are also plotted in figure 23 and are in good agreement with experiment.

The magnitude of the interaction of μ mesons with nuclei has been deduced from these data (see Sard and Crouch) and is of the same order as the interaction responsible for β decay.

6.2. π *Mesons : Level Width Measurements*

The interaction with the nucleus in the case of π mesons is at least 10^{12} times stronger than for μ mesons and the radioactive decay of the π meson (mean life 2×10^{-8} sec) cannot compete with capture. The interaction is so rapid however that appreciable broadening of the energy levels occurs.

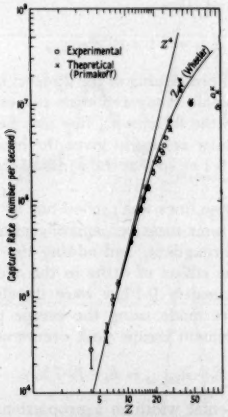


Figure 23. Rate of capture of μ^- mesons from the 1s level. The experimental values are those of Sens *et al.* (1957).

Brueckner (1955) has calculated that the broadening of the 1s level due to the absorption process involving two nucleons is given by

$$\Gamma_{1s} = \frac{8}{3}\frac{Z^2}{2150} E_{K_\alpha} \qquad \ldots\ldots(39)$$

where Γ_{1s} is the full width at half maximum for the 1s level. Equation (39) is derived from the capture rate in deuterium deduced from the probability of the reaction $(p + p \rightarrow \pi^+ + D)$ (Brueckner *et al.* 1951).

The level broadening has been measured in the case of Be, where Γ is expected to be 0·9 kev according to equation (39), by West and Bradley (1957). They compared the widths of the π mesonic fluorine $L_\alpha(3d \rightarrow 2p)$ line (energy 41·7 kev) and the beryllium $K_\alpha(2p \rightarrow 1s)$ line (energy 42·1 kev) using a proportional counter.

The natural broadening of the 2p levels involved in these transitions can be estimated from the parameters obtained from the total K yields (equations (32), (33) and (34)).

We have

$$\Gamma_{2p} \simeq \hbar\omega_{\text{total}} = \hbar(\omega_{\text{coll}} + \omega_{\text{Auger}} + \omega_r + \omega_{\text{abs}}) \qquad \ldots\ldots(40)$$

and for fluorine

$$\omega_{\text{coll}} + \omega_{\text{Auger}} \qquad = 4 \quad \times 10^{13}\ \text{sec}^{-1}$$

$$\omega_r \qquad = 1{\cdot}75 \times 10^{11}\, Z^4 = 1{\cdot}2 \times 10^{15}\ \text{sec}^{-1}$$

$$\omega_{\text{abs}} \qquad = 3{\cdot}5 \times 10^{9}\, Z^7 \quad = 1{\cdot}7 \times 10^{16}\ \text{sec}^{-1}.$$

The major part of the transition rate from the 2p state in fluorine therefore arises from the absorption term and since

$$\Gamma_{(ev)} = 6{\cdot}6 \times 10^{-16}\, \omega_{\text{total}} \qquad \ldots\ldots(41)$$

one deduces that the natural broadening of the 2p level in fluorine is approximately 10 ev. This is quite negligible compared with the expected broadening of the 1s level of beryllium and so the fluorine L_α line may be used for determining the instrumental width. A similar argument gives the broadening of the 2p level in beryllium as approximately 0·1 ev so that the width of the $2p \rightarrow 1s$ line is essentially that of the 1s level.

The comparison of the two lines was carried out by alternating runs with LiF and beryllium metal of the same mass per square centimetre, using approximately the same beam intensity throughout, and adding the individual sets of spectra at the end of a series. The effects of drifts in the overall gain of the apparatus which amounted to approximately 0·1 kev were thereby minimized. The most accurate measurements were made using the escape peaks in the proportional counter spectra. The prominent escape peak occurs at an energy

$$E_1 - E_{eK_\alpha Xe} = E_1 - 29{\cdot}7\ \text{kev}.$$

Since the absolute instrumental width in a proportional counter is proportional to $E^{1/2}$, the magnitude of the instrumental width is much smaller for the escape peaks and more precise measurements of any extra width in the case of the beryllium line can be made. Figure 24 shows two spectra of the escape peaks from the beryllium K_α line and the fluorine L_α line. The widths, measured by fitting Gaussian distributions to the spectra, then give the natural broadening of the beryllium line from the relation

$$\Gamma_{\text{nat}}^2 = \Gamma_{\text{Be}}^2 - c\Gamma_{\text{F}}^2 \qquad \ldots\ldots(42)$$

where the factor c takes account of the slight difference between the energies of the two lines.

A mean of three measurements gave

$$\Gamma_{\text{nat}} = 1{\cdot}2 \pm 0{\cdot}2\ \text{kev} \qquad \ldots\ldots(43)$$

which is very close to the theoretical estimate (0·9 kev) given by Brueckner. The experimental result corresponds to a capture rate from the 1s level of beryllium

$$\omega_{\text{abs}\,1s} = 1{\cdot}8 \times 10^{18}\ \text{sec}^{-1} \qquad \ldots\ldots(44)$$

and corresponds to an interaction rate for π mesons with nuclei some 3×10^{13} times that for μ mesons (cf. figure 23).

The instrumental width of the scintillation counter is too broad to allow the same technique to be used for measuring line widths. However the partial absorption of a line, by a critical absorber whose edge falls within the natural line width, has enabled an estimate of the width to be made in the case of the

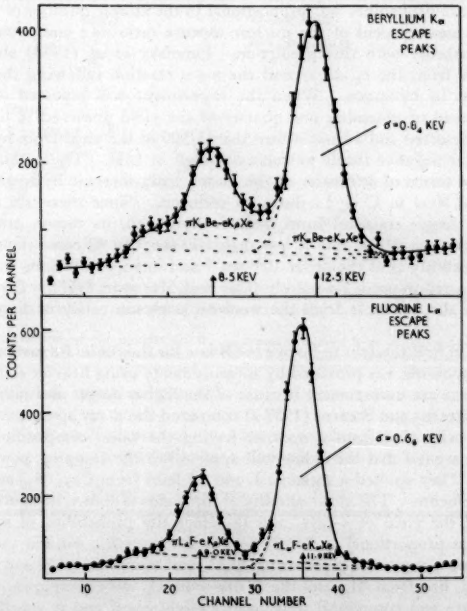


Figure 24. Comparison of the widths of the escape peaks from Be K_α and F L_α radiations. (*After* West and Bradley 1957.)

^{10}B K_α line. Stearns and Stearns (1956) by this means obtained a width $\Gamma_B = 2{\cdot}2 \pm 2{\cdot}8$ kev consistent with the value 2·2 kev given by Brueckner's relation (39).

It was pointed out by Fitch and Rainwater that appreciable broadening of the 2p levels in μ mesonic atoms is to be expected at high Z due to the high radiative transition probability. For lead $(\omega_r)_{2p} \simeq 10^{18}$ sec^{-1} giving $\Gamma_{2p} \simeq 1$ kev. However the broadening would be difficult to detect as even the L-ray energies are about 2·6 Mev for lead.

§ 7. Relative Probability of the Capture of a Meson in the Constituents of a Chemical Compound

Fermi and Teller (1947) estimated the relative probability of mesonic atom formation in the constituents of a compound. They assumed that the probability of formation of a mesonic atom was proportional to the rate of energy loss of the meson at low energy in the neighbourhood of the particular atom, and deduced that the capture probability was proportional to the atomic number of the element. The earliest measurement of the nuclear capture ratio in a compound, however, disagreed markedly with this prediction. Panofsky *et al.* (1951) observed high energy γ-rays from the π_0 decay and the $n+\gamma$ reaction following the absorption of π^- mesons in hydrogen. When the experiment was repeated with $(CH_2)_n$, however, instead of obtaining one quarter of the yield observed in hydrogen, no γ-rays were detected and a limit of less than 1/500 of the yield from hydrogen was set. A similar negative result was also obtained in LiH. The authors explained their result in terms of a transfer of the meson from mesonic hydrogen to a more tightly bound level in C or Li during a collision. Since there are no electrons available for Auger emission from mesonic hydrogen, its meson can only make radiative transitions whose rate is less than 10^{11} sec^{-1} in all cases (figure 4). The collision probability is of the order 10^{13} sec^{-1} and in such collisions transfer of the meson is expected to occur frequently (Day and Morrison 1957). Direct evidence of transfer is also available from the work on μ meson catalysis discussed in the next section.

Evidence at first thought to favour the Z law for the *initial* formation of mesonic atoms in compounds was provided by measurements using heavier elements where exchange effects are unimportant because of the higher Auger and radiative transition rates. Stearns and Stearns (1957 a) compared the x-ray spectrum obtained in a compound with that from a mixture having the same composition. For the mixture they argued that the Z law will apply since the stopping power per atom varies as Z. They studied π mesonic L and M lines from Ca_{20}, S_{16} and a mixture of the two elements. The close equality of the yields is shown in figure 25. They deduced that the yield of x-rays, and therefore the probability of mesonic atom formation, was proportional to $Z^{1\cdot0\pm0\cdot2}$ in this compound. Similar measurements with μ mesonic x-rays were made using Al_2O_3 and a mixture of Al and H_2O. The yields of the L line from Al_{13} and the K line from O_8 were compared. The yields for the mixture and compound were again closely equal and it was deduced that the probability of mesonic atom formation in the compounds was proportional to $Z^{1\cdot0\pm0\cdot1}$.

These conclusions have however been criticized by Sens *et al.* (1958), who pointed out that estimates of the atomic stopping power for low energy mesons, obtained from the range energy relations, are not in accordance with an atomic stopping power proportional to Z. In fact even for 1 Mev π mesons (Rich and Madey 1954), for instance, the atomic stopping power of O is nearly four times that of Al. The ratio of the atomic stopping powers is not known accurately at very low energies and Stearns' results, therefore, only show that there is no effect of chemical binding on the ratio of mesonic atom formation in the constituents of CaS and Al_2O_3.

Direct evidence on the capture ratio in the constituents of compounds was recently obtained by Sens *et al.* (1958). They measured the decay curves of μ^- mesons stopped in compounds. From the composite decay curves obtained, knowing the characteristic decay lifetimes for μ^- mesons in the individual constituents, they were able to deduce the relative numbers of mesons reaching the 1s levels in the constituents of the compounds.

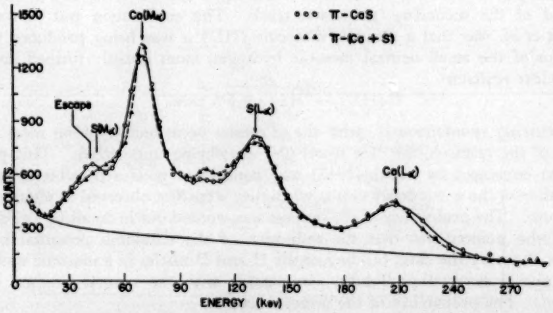


Figure 25. Comparison of the spectra of the π mesonic x-rays from a compound (CaS) and a mixture containing the same proportions of Ca and S. (*After* Stearns and Stearns 1957 a.)

Their results are shown in table 7.

Table 7. Relative Numbers of μ^- Mesons reaching the 1s Level in the Constituents of a Compound

Compound	Ratio	Observed Sens *et al.*	Predicted Fermi & Teller	Atomic ratio
P_2O_5	O/P	2·7±0·3	1·3	2·5
Cl_4	Cl/C	4·1±0·8	11·3	4·0
Al_2O_3	O/Al	2·3±0·2	0·92	1·5
SiO_2	O/Si	3·5±0·3	1·1	2·0
KHF_2	F/K	1·7±0·4	0·95	2·0
KOH	O/K	2·2±0·4	0·42	1·0
$C_6H_4Cl_2$ (Liquid)	C/Cl	2·3±0·2	1·06	3·0
$C_6H_4Cl_2$ (Solid)	C/Cl	2·1±0·2	1·06	3·0

The observed ratios agree much better with capture occurring in proportion to the number of atoms of each constituent than with the Fermi–Teller Z law.

Moreover, any exchange of mesons from the lighter to the heavier constituent, which might have occurred before the 1s level was reached, would make the disagreement with the Z law for the initial formation of the mesonic atom even more pronounced.

§ 8. Mesonic Molecules and the Catalysis of Nuclear Reactions between Hydrogen Isotopes

It was observed by Alvarez, Bradner, Crawford *et al.* (1957) that, when μ^- mesons were brought to rest in a liquid hydrogen bubble chamber, the μ^- meson was occasionally re-emitted with an energy of 5·3 Mev. Moreover there was often an appreciable gap between the end of the incident track and the start of the secondary track. On much rarer occasions a similar event also occurred at the end of the secondary μ^--meson track. The explanation put forward by Alvarez *et al.* was that a mesonic molecule $(HD)^+\mu$ was being produced during collisions of the small neutral mesonic hydrogen atom initially formed and that the nuclear reaction

$$(HD)_\mu \rightarrow {}^3He + \gamma + 5{\cdot}5 \text{ Mev} \qquad \ldots\ldots(45)$$

was occurring spontaneously, with the μ meson occasionally taking most of the energy of the reaction, the ^{3}He recoil (0·2 Mev) being undetected. This process was first envisaged by Frank (1947) who considered it as a possible alternative explanation of the $\pi \rightarrow \mu$ decay events when they were first observed in photographic emulsions. The probability of the process was worked out in detail by Zel'dovitch (1954) who pointed out that the reduction of the Coulomb potential barrier, brought about by the close binding of the H and D nuclei in a mesonic molecule, would give a reaction probability comparable with the natural lifetime of the μ meson. The probability of the direct reactions

$$H_\mu + D \text{ or } D_\mu + H \rightarrow {}^3He + \gamma \qquad \ldots\ldots(46)$$

is too small (reaction rate approximately 1 sec^{-1} even neglecting the Coulomb barrier completely (Skyrme 1957)) to compete with natural decay and therefore cannot account for the reactions observed.

The latest data on the observed frequency of re-emission of the μ meson have been given by Cresti and Gottstein (1957, private communication) and are reproduced in table 8.

Table 8. Frequency of Catalysed (H, D) Reactions with Re-emission of the μ Meson in Liquid Hydrogen

Deuterium concentration (%)	0·016 (natural)	0·3	4·3
Number of events $\mu^- \rightarrow e^- + 2\nu$	2547	5924	1792
Number of events $(HD)_\mu \rightarrow {}^3He + \mu$	18	139	45
Number of double events $(HD)_\mu \rightarrow {}^3He + \mu$ with μ re-emitted on 2 successive occasions	0	3	0
Ratio $({}^3He + \mu)$ reactions per μ^- stopped (%)	0·7 ± 0·2	2·3 ± 0·2	2·5 ± 0·4

The nuclear reaction frequency, of which only the part involving re-emission of the μ meson is observed directly, is much too high and saturates at too low a concentration to be accounted for by molecule formation as a result of the process

$$H_\mu + D \rightarrow (HD)_\mu^*. \qquad \ldots\ldots(47)$$

It is necessary to suppose that a transfer of the μ meson first takes place from H_μ to D_μ, in which the meson is more tightly bound as a result of the difference

in the reduced masses,† and that the $(HD)_\mu$ is formed in a subsequent collision with hydrogen.

Skyrme has worked out the probability of the re-emission of the μ meson phenomenologically in terms of the scheme shown in figure 26. The λ's are reaction rates and λ_0 is the rate of decay of the μ meson ($4{\cdot}5 \times 10^5$ sec^{-1}). He obtains, at low concentrations c of D,

$$1/Y = A + B/c. \qquad \ldots\ldots(48)$$

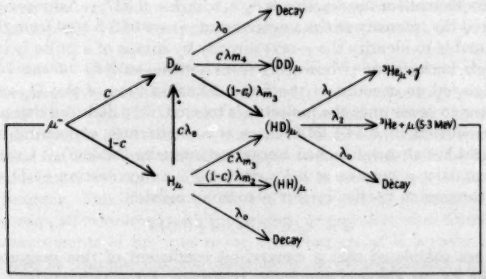


Figure 26. Processes involved in μ meson catalysed fusion reactions in liquid hydrogen. c is the concentration of deuterium.

Y is the total yield of reactions in which the μ meson is re-emitted (including multiple reactions) where

$$A+1 = \frac{(\lambda_0+\lambda_{m_2})}{\lambda_{m_2}}\frac{(\lambda_0+\lambda_1+\lambda_2)}{\lambda_2}, \quad B = (A+1)\frac{(\lambda_0+\lambda_{m_3})}{\lambda_0}$$

provided $\lambda_e \gg \lambda_0$ and λ_{m_2} and λ_{m_3} and the symbols refer to the processes indicated in figure 26. This relation fits the data in table 7 with $A = 39$ and $B = 0{\cdot}016$, giving

$$\lambda_0/(\lambda_0+\lambda_{m_3}) = 2500. \qquad \ldots\ldots(49)$$

The probability of a meson transferring from H to D is, at low concentration of D, given by

$$p_t = \frac{c\lambda_e}{c\lambda_e+\lambda_0+\lambda_{m_3}} \qquad \ldots\ldots(50)$$

where λ_{m_3} has been equated to λ_{m_4}. Cresti and Gottstein were able to derive this probability independently from their observations of the frequency of gaps at the ends of μ-meson tracks. They observed gaps of mean length 0·6 mm between the primary and secondary μ-meson tracks and also between μ-meson and decay electron tracks. The size of the gaps is too large to be accounted for by diffusion of the H_μ and they must arise as a result of the recoil energy (43 ev) imparted to the D_μ during the exchange of the meson; a similar recoil will also occur during

† Skyrme has estimated the following binding energies (kev): H_μ, 2·542; $(HD)_\mu$, 2·941; D_μ, 2·677; $(DD)_\mu$, 3·038; $(HH)_\mu$, 2·824.

the molecule formation. By comparing the frequency of gaps when the secondary was a μ meson, and therefore the transfer had certainly taken place, with the corresponding frequency when the secondary was an electron, they were able to deduce the total probability of transfer. They obtained, at 0·016% concentration, $p_t = 0{\cdot}55 \pm 0{\cdot}29$ and at 0·3% concentration $p_t = 0{\cdot}76 \pm 0{\cdot}16$. The corresponding quantities deduced from equations (49) and (50) are 0·28 and 0·88, in satisfactory agreement.

From the expression for $A+1$ in equation (48) it can be shown that the internal conversion coefficient for the reaction is $\lambda_2/\lambda_1 > 1/A = 0{\cdot}027$. Ashmore *et al.* (1958) have measured the intensity of the unconverted γ-rays of 5·5 MeV from the reaction. They were unable to identify the γ-rays directly by means of a pulse height analysis due to a high background; however, from a time analysis of the γ-ray pulses, which are delayed on account of the finite diffusion time of the D_μ and the time for the reaction to occur once the molecule is formed, they deduced that unconverted γ-rays were emitted in $0{\cdot}34 \pm 0{\cdot}06$ cases at a deuterium concentration of 1·8% where the yield has already reached saturation intensity (table 8). Using the mean yields of secondary μ mesons at 0·3% and 4·3% concentration of D one obtains an internal conversion coefficient for μ-meson emission

$$\alpha = \lambda_2/\lambda_1 = 0{\cdot}069 \pm 0{\cdot}014. \qquad \ldots\ldots(51)$$

Skyrme has calculated that a conversion coefficient of this magnitude can be accounted for by an electric monopole transition for which $\alpha \sim 0{\cdot}17$, this value being probably an overestimate due to the approximations made.

When a γ-ray is emitted, the μ meson probably remains bound to the ^{3}He nucleus and is lost for further reactions.† Considering only subsequent reactions produced by the 5·3 MeV meson, it can be shown that the ratio of the number of events in which a μ meson catalyses more than one reaction to the number of events in which it catalyses one reaction is equal to Y. At saturation this ratio is therefore equal to 0·024. From the observed number of single catalysis events in table 8, the expected total number of multiple events is 4·5. The observed number of double events, 3, is therefore in agreement with this estimate. An accurate measurement of the frequency of multiple events would show whether the μ mesons bound to the ^{3}He nucleus were detached often enough to affect the yield of multiple reactions.

Other reactions between hydrogen isotopes have been considered by Zel'dovitch and by Jackson. The calculated nuclear reaction rates in the mesonic molecule are given in table 9.

Table 9. Calculated Nuclear Reaction Rates in Mesonic Molecules

Molecule	$(HD)_\mu$	$(DD)_\mu$	$(TD)_\mu$
Reaction rate (sec^{-1})	$0{\cdot}3 \times 10^8$	$0{\cdot}7 \times 10^{11}$	$0{\cdot}4 \times 10^{13}$

The D + D and T + D reactions will therefore compete much more favourably with the natural decay of the μ meson, once the molecule is formed. While there is no chance of the μ meson taking any of the reaction energy in these cases, the

† Jackson (1957) gives an expression for the probability of the μ meson remaining bound to a moving particle. In the case of the slow ^{3}He nucleus produced in the (pD) reaction this probability is estimated at 98%.

charged reaction products are much more energetic than for the p+D reaction and there is an appreciable chance of the μ meson becoming detached from inert atoms. Jackson has estimated that the probability of the meson remaining bound to an inert atom is only 8% in the D+D reaction and 1% in the T+D reaction. However the natural decay of the μ meson during the time taken to form the mesonic molecule will reduce the chance of subsequent reactions occurring and it is not possible to give a reliable estimate of the number of reactions to be expected. Except for rare D+D reactions observed in liquid hydrogen, the D+D and T+D reactions have not yet been studied. For π mesons nuclear capture from the 1s level is so rapid ($\sim 10^{-15}$ sec) that catalysis cannot occur.

§ 9. Mesonic Atoms formed with K Mesons and Hyperons

The other negative particles which have been produced in high energy reactions (K^-, anti-proton, Σ^- and Ξ^-) would also be expected to form 'mesonic' atoms. The heaviest hyperon (Ξ^-) will, however, have its K orbit radius smaller than the nuclear radius for $Z>7$ and only the higher energy levels will be hydrogen-like.

K^- mesons (mass $966\,m_e$) have been observed to come to rest in liquid hydrogen (Alvarez, Bradner, Falk-Vairant *et al.* 1957) and out of a total of 126 mesons observed to stop, all interacted with the proton, presumably after forming a mesonic atom. Measurements of the interaction cross section of K mesons at low energy show that the K^- meson–nucleon interaction is of comparable strength to the π meson–nucleon interaction. If we assume that the interaction with nucleons is the same for K^- mesons and π^- mesons, the higher mass of the K meson would have the following consequences for the K mesonic atom.

(i) Binding energies of the levels would be $\bar{m}_K/\bar{m}_\pi \simeq 3{\cdot}54$ times the π mesonic energies neglecting nuclear interactions.

(ii) K series x-rays.

The probability of absorption from the 2p state (appendix, eqn (A5)) is proportional to $\bar{\mu}^5$. The radiative probability increases as $\bar{\mu}$, and assuming that $\omega_{\text{coll}} = 2\times 10^{13}$ sec^{-1}, one can scale equations (33) and (34) for π mesons to derive a K x-ray yield for K mesonic atoms. The x-ray yield derived in this way has a maximum value at $Z=2$ of the order of 4×10^{-2} and falls to approximately 2×10^{-3} at $Z=6$. A yield as low as this, combined with the rather low intensity of K mesons produced by present machines, would make the measurement of K series x-rays extremely difficult. In addition absorption from the 1s level varies as $\bar{\mu}^3$ (appendix, eqn (A4)) so that the 1s level broadening in Be for example would be about 50 kev, the energy of the line being about 160 kev.

(iii) L series x-rays.

Absorption from the 3d state (appendix, eqn (A6)) varies as $\bar{\mu}^7$. Scaling equations (29) and (31) and using $\omega_{\text{coll}} = 10^{14}$ sec^{-1} the calculated L series yield has a maximum value of 0·6 at $Z=5$ falling to 10^{-2} at $Z=20$. The L x-rays should therefore be measurable at low atomic numbers. It should be emphasized that these predictions about yields are very tentative and are given merely to illustrate the trends to be expected in the heavy mesonic atoms.

Finally Auger electrons have been observed at the end of Σ^- (mass $2341\,m_e$) hyperon tracks in photographic emulsions by Fry *et al.* (1957). The yield appears to be comparable with that obtained from μ mesons.

ACKNOWLEDGMENTS

I am indebted to Dr. M. B. Stearns for preprints of the work on π mesonic x-ray yields and for making available her own review article (Stearns 1957) before publication. Other recent reviews of the subject have been published by DeBenedetti (1956) and Ivanenko and Pustovalov (1957).

I am indebted to Dr. J. S. Bell for clarification of several theoretical points, also to Dr. E. H. S. Burhop for information on the Auger yields and to Professor R. Karplus for elucidating the theoretical interpretation of the π mesonic x-ray shifts.

I also wish to thank the authors and publishers of the papers from which many of the diagrams have been taken, for permission to reproduce them.

APPENDIX

THE OVERLAP OF SCHROEDINGER WAVE FUNCTIONS WITH THE NUCLEUS

The Schroedinger wave function in a hydrogen-like atom (Condon and Shortley 1935) is

$$\psi_{ni} = \frac{1}{2\sqrt{\pi}} \left\{ \frac{4Z^3}{n^4 a^3} \frac{(n-l-1)!}{[(n+l)!]^3} \right\}^{1/2} e^{-\rho/2} \rho^l L_{n+l}^{2l+1}(\rho) \qquad \ldots\ldots(A1)$$

which is normalized so that

$$\int_0^\infty 4\pi r^2 \psi_{ni}^2 \, dr = 1$$

and

$$\rho = 2Zr/na \text{ with } a = \text{Bohr radius of the bound particle} = \hbar^2/\tilde{\mu}e^2.$$

$L_{n+l}^{2l+1}(\rho)$ takes the following values

$1s: -1!$	$2s: 2\rho - 4$
$2p: -3!$	$3s: -3\rho^2 + 18\rho - 18$
$3d: -5!$	$3p: 24\rho - 96$
$4f: -7!$	$4d: 720\rho - 5760$

The constant term in the expression for L_{n+l}^{2l+1} is in general

$$\frac{-\{(n+l)!\}^2}{(n-l-1)!\,(2l+1)!}$$

consequently for s states at $r = 0$

$$\psi_{n,0} = -\frac{1}{2\sqrt{\pi}} \left\{ \frac{4Z^3}{n^4 a^3} \frac{(n+l)!}{(n-l-1)!\,(2l+1)!^2} \right\}^{1/2}$$

$$= -\frac{1}{2\sqrt{\pi}} \left\{ \frac{4Z^3}{n^3 a^3} \right\}^{1/2}.$$

Hence

$$|\psi_{r=0}|^2 = \frac{1}{\pi} \frac{Z^3}{n^3 a^3}. \qquad \ldots\ldots(A2)$$

The probability of capture by the nucleus can be written

$$P_{\text{abs}} \propto \int_0^R q_r \, 4\pi r^2 \psi_{\text{mi}}{}^2 \, dr$$

where q_r = proton density within the nucleus at $r = Z/(4\pi R^3/3)$. For the 1s state, the probability of the meson being inside the nucleus is

$$\int_0^R 4\pi r^2 \psi_{1,0}{}^2 \, dr = \frac{1}{2}\int_0^{2ZR/a} e^{-\rho} \rho^2 \, d\rho$$

$$= \tfrac{1}{2}\left|\frac{\rho^3}{3} - \frac{\rho^4}{4} + \ldots\right|_0^{2ZR/a} \qquad \ldots\ldots\text{(A3)}$$

for $2ZR/a$ small $\simeq \frac{1}{6}\left(\frac{2ZR}{a}\right)^3$.

Hence

$$(P_{\text{abs}})_{1s} \propto \frac{3Z}{4\pi R^3}\frac{1}{6}\left(\frac{2ZR}{a}\right)^3$$

$$\propto \frac{Z^4}{\pi a^3} \propto \bar{\mu}^3 Z^4. \qquad \ldots\ldots\text{(A4)}$$

This expression is equal to $|\psi_0|^2 Z$ and so neglecting ρ^4 and higher terms in (A3) is equivalent to taking the 1s wave function as constant over the nucleus. For $l > 0, |\psi_0|^2$ is zero and this result cannot be used for higher l values.

For the 2p state

$$\int_0^R 4\pi r^2 \psi_{2,1}{}^2 \, dr = \frac{1}{24}\int_0^{ZR/a} \rho^4 e^{-\rho} \, d\rho$$

$$= \frac{1}{24}\left|\frac{\rho^5}{5} - \frac{\rho^6}{6} + \ldots\right|_0^{ZR/a}$$

$$= \frac{1}{120}\left(\frac{ZR}{a}\right)^5$$

neglecting terms in ρ^6 and higher.

Hence

$$(P_{\text{abs}})_{2p} \propto \frac{Z}{4\pi R^3/3}\left(\frac{ZR}{a}\right)^5$$

$$\propto \frac{Z^6 R^2}{a^5};$$

since $R = r_0 A^{1/3} \propto Z^{1/3}$ approximately

$$(P_{\text{abs}})_{2p} \propto \frac{Z^{6.7}}{a^5} \propto \bar{\mu}^5 Z^{6.7}. \qquad \ldots\ldots\text{(A5)}$$

The relation, often quoted, that $(P_{\text{abs}})_{2p} \propto Z^6$ is derived by evaluating $|\psi|^2$ at the nuclear surface. This varies as Z^{2l+3} if the Z dependence of the nuclear radius is neglected and multiplying by Z, the number of protons in the nucleus, one obtains an absorption probability proportional to Z^{2l+4}.

In a similar way to that used in deriving (A4) and (A5) it can be shown that

$$(P_{\text{abs}})_{3d} \propto \frac{Z^8}{a^7} R^4 \propto \bar{\mu}^7 Z^{b-3}. \qquad \ldots\ldots(A6)$$

The general expression, to the same degree of accuracy, can be derived using only the constant term in $L_{n+l}^{2l+1}(\rho)$.

Then

$$\int_0^R 4\pi r^2 \psi_{nl}{}^2 \, dr = \left(\frac{2Z}{na}\right)^{2l} \frac{4Z^3}{n^4 a^3} \frac{(n+l)!}{(n-l-1)!\,(2l+1)!^2} \frac{R^{2l+3}}{2l+3}. \qquad \ldots\ldots(A7)$$

Hence

$$(P_{\text{abs}})_{(n,\,l)} \propto \frac{Z}{R^3} \int_0^R 4\pi r^2 \psi_{nl}{}^2 \, dr$$

$$\propto \frac{Z^{8l/3+4}}{a^{2l+3}} \frac{(n+l)!}{n^{2l+4}(n-l-1)!\,(2l+1)!^2(2l+3)} \qquad \ldots\ldots(A8)$$

and for $l = n-1$

$$P_{\text{abs}} \propto \frac{Z^{8l/3+4}}{a^{2l+3}} \frac{1}{n^{2n+2}(2n-1)!\,(2n+1)}. \qquad \ldots\ldots(A9)$$

References

Alvarez, L. W., Bradner, H., Crawford, F. S., Jr., Crawford, J. A., Falk-Vairant, P., Good, N. L., Gow, J. D., Rosenfeld, A. H., Solmitz, F. T., Stevenson, M. L., Tichо, H. K., and Tripp, R. D., 1957, *Phys. Rev.*, **105**, 1127.

Alvarez, L. W., Bradner, H., Falk-Vairant, P., Gow, J. D., Rosenfeld, A. H., Solmitz, F. T., and Tripp, R. D., 1957, *Nuovo Cim.*, **5**, 1026.

Ashmore, A., Nordhagen, R., Strauch, K., and Townes, B. M., 1958, *Proc. Phys. Soc.*, **71**, 161.

Barkas, W. H., Birnbaum, W., and Smith, F. M., 1956, *Phys. Rev.*, **101**, 778.

Bethe, H. A., and de Hoffman, F., 1955, *Mesons and Fields*, Vol. II (Evanston, Illinois : Row, Peterson).

Bethe, H. A., and Salpeter, E. E., 1957, *Handbuch der Physik*, Vol. 35 (Berlin : Springer).

de Borde, A. H., 1954, *Proc. Phys. Soc.* A, **67**, 57.

Brueckner, K. A., 1955, *Phys. Rev.*, **98**, 769.

Brueckner, K. A., Serber, R., and Watson, K., 1951, *Phys. Rev.*, **81**, 575.

Burbidge, G. R., and de Borde, A. H., 1953, *Phys. Rev.*, **89**, 189.

Butement, F. D. S., 1953, *Phil. Mag.*, **44**, 208 ; 1955, *Ibid.*, **46**, 1136.

Camac, M., Halbert, M. L., and Platt, J. B., 1955, *Phys. Rev.*, **99**, 905.

Camac, M., McGuire, A. D., Platt, J. B., and Schulte, H. J., 1952, *Phys. Rev.*, **88**, 134 ; 1955, *Ibid.*, **99**, 897.

Cassels, J. M., O'Keeffe, T. W., Rigby, M., Wetherell, A. M., and Wormald, J. R., 1957, *Proc. Phys. Soc.* A, **70**, 543.

Chang, W. Y., 1949, *Rev. Mod. Phys.*, **21**, 166 ; 1954, *Phys. Rev.*, **95**, 1288.

Cohen, E. R., Crowe, K. M., and DuMond, J. W. M., 1956, *Phys. Rev.*, **104**, 266.

Condon, E. U., and Shortley, G. H., 1935, *The Theory of Atomic Spectra* (Cambridge : University Press).

Conversi, M., Pancini, E., and Piccioni, O., 1947, *Phys. Rev.*, **71**, 209.

Cooper, L. N., and Henley, E. M., 1953, *Phys. Rev.*, **92**, 801.

Cosyns, M. G. E., Dilworth, C. C., Occhialini, G. P. S., Schoenberg, M., and Page, N., 1949, *Proc. Phys. Soc.* A, **62**, 801.

Crowe, K. M., 1957, *Nuovo Cim.*, **5**, 541.

Day, T. B., and Morrison, P., 1957, *Phys. Rev.*, **107**, 912.

DeBenedetti, S., 1956, *Nuovo Cim.*, **4**, Suppl. 3, 1234.

DEMEUR, M., 1956, *Nuclear Physics*, **1**, 516.
DESER, S., GOLDBERGER, M. L., BAUMANN, K., and THIRRING, W., 1954, *Phys. Rev.*, **96**, 774.
FAFARMAN, A., and SHAMOS, M. H., 1955, *Phys. Rev.*, **100**, 874.
FERMI, E., and TELLER, E., 1947, *Phys. Rev.*, **72**, 399.
FITCH, V. L., and RAINWATER, J., 1953, *Phys. Rev.*, **92**, 789.
FOLDY, L. L., 1954, *Phys. Rev.*, **93**, 880.
FOLDY, L. L., and ERIKSEN, E., 1954, *Phys. Rev.*, **95**, 1048.
FRANK, F. C., 1947, *Nature, Lond.*, **160**, 525.
FRY, W. F., 1953, *Nuovo Cim.*, **10**, 490.
FRY, W. F., SCHNEPS, J., SNOW, G. A., SWAMI, M. S., and WOLD, D. C., 1957, *Phys. Rev.*, **107**, 257.
GARWIN, R. L., LEDERMAN, L. M., and WEINRICH, M., 1957, *Phys. Rev.*, **105**, 1415.
HILL, D. L., and FORD, K. W., 1954, *Phys. Rev.*, **94**, 1617.
IVANENKO, D. D., and PUSTOVALOV, G. E., 1957, *Uspekhi, Fiz. Nauk.*, **61**, 27.
JACKSON, J. D., 1957, *Phys. Rev.*, **106**, 330.
JACOBSOHN, B. A., 1954, *Phys. Rev.*, **96**, 1637.
KARPLUS, R., 1956, *Proceedings of the Sixth Annual Rochester Conference* (New York: Interscience).
KARPLUS, R., and HALPERN, F., 1957, *Bull. Amer. Phys. Soc.*, **2**, 5.
KOSLOV, S., FITCH, V. L., and RAINWATER, J., 1954, *Phys. Rev.*, **95**, 291.
LAKIN, W., and KOHN, W., 1954, *Phys. Rev.*, **94**, 787.
LAMB, W. E., and RETHERFORD, R. C., 1950, *Phys. Rev.*, **79**, 549.
MARSHAK, R. E., 1952, *Meson Physics* (New York: McGraw Hill).
MENON, M. G. K., MUIRHEAD, H., and ROCHAT, O., 1950, *Phil. Mag.*, **41**, 583.
MESSIAH, A. M. L., and MARSHAK, R. E., 1952, *Phys. Rev.*, **88**, 678.
MICKELWAIT, A. B., and CORBEN, H. C., 1954, *Phys. Rev.*, **96**, 1145.
OREAR, J., 1956, *Nuovo Cim.*, **4**, 856.
PANOFSKY, W. K. H., AAMODT, R. L., and HADLEY, J., 1951, *Phys. Rev.*, **81**, 565.
PRIMAKOFF, H., 1955, *Proceedings of the Fifth Annual Rochester Conference* (New York: Interscience).
RICH, M., and MADCY, R., 1954, *U.C.R.L.* 2031.
SARD, R. D., and CROUCH, M. F., 1954, *Progress in Cosmic Ray Physics*, Vol. II, ed. J. G. Wilson (Amsterdam: North-Holland).
SCOTT, J. M. C., 1956, *Progress in Nuclear Physics*, Vol. 5, ed. O. R. Frisch (London: Pergamon Press).
SENS, J. C., SWANSON, R. A., TELEGDI, V. L., and YOVANOVITCH, D. D., 1957, *Phys. Rev.*, **107**, 1464.
SKYRME, T. H. R., 1957, *Phil. Mag.*, **2**, 910.
STEARNS, M., and STEARNS, M. B., 1956, *Phys. Rev.*, **103**, 1534; 1957 b, *Ibid.*, **107**, 1709.
STEARNS, M., STEARNS, M. B., LEIPUNER, L., and DEBENEDETTI, S., 1954, *Phys. Rev.*, **95**, 625.
STEARNS, M. B., 1957, *Progress in Nuclear Physics*, Vol. 6, ed. O. R. Frisch (London: Pergamon Press).
STEARNS, M. B., and STEARNS, M., 1957 a, *Phys. Rev.*, **105**, 1573.
STEARNS, M. B., STEARNS, M., DEBENEDETTI, S., and LEIPUNER, L., 1954 a, *Phys. Rev.*, **95**, 1353; 1954 b, *Ibid.*, **96**, 804.
STEARNS, M. B., STEARNS, M., and LEIPUNER, L., 1957, *Phys. Rev.*, **108**, 445.
TICHO, H. K., 1948, *Phys. Rev.*, **74**, 1337.
UEHLING, E. A., 1935, *Phys. Rev.*, **48**, 55.
WEST, D., BATCHELOR, R., and BRADLEY, E. F., 1955, *Proc. Phys. Soc.* A, **68**, 801.
WEST, D., and BRADLEY, E. F., 1956, *Phil. Mag.* (Ser. 8), **1**, 97; 1957, *Ibid.*, **2**, 957.
WHEELER, J. A., 1947, *Phys. Rev.*, **71**, 320; 1949, *Rev. Mod. Phys.*, **21**, 133; 1953, *Phys. Rev.*, **92**, 812.
WILETS, L., 1954, *K. Danske Mat.-Fys. Medd.*, **29**, no. 3.
WOLFENSTEIN, L., 1957, *Bull. Amer. Phys. Soc.*, **2**, 39.
ZEL'DOVITCH, YA. B., 1954, *Doklady Akad. Nauk S.S.S.R.*, **95**, 493.

PRODUCTION AND MEASUREMENT OF HIGH TEMPERATURES

By W. LOCHTE-HOLTGREVEN
Kiel University

With an Addendum (§ 6) by
Dr. R. SCHALL and Dr. F. WECKEN
Weil/Rhein

CONTENTS

PAGE

§ 1. Measurement of high temperatures ... 314
1.1. Introduction ... 314
1.1.1. Temperature and temperature equilibrium ... 314
1.1.2. Time required to establish temperature equilibrium 315
1.1.3. Effect of a temperature gradient in a light source .. 317
1.1.4. Effect of the optical depth of the emitting gas ... 317
1.2. Determination of temperature ... 317
1.2.1. The determination of the gas temperature in the case of an optically thick layer ... 317
1.2.2. The determination of the gas temperature in the case of an optically thin layer ... 318
1.2.3. The determination of the excitation and ionization temperature ... 319
1.3. The use of line broadening for the determination of the density n_e of electrons ... 321
1.3.1. Line broadening by impact ... 321
1.3.2. Statistical line broadening ... 322
1.3.2. (*a*). Perturbation by one particle ... 322
1.3.2. (*b*). Perturbation by several particles ... 323
1.3.3. Experimental proof of Holtsmark's theory ... 324
1.3.4. The evaluation of n_e from the apparent series limit 326
1.3.5. The diminution of the ionization potential within a plasma ... 328
1.3.6. The continuous radiation beyond the series limit .. 328
1.4. Determination of temperatures beyond 15 000°K ... 330
1.4.1. Temperature evaluation from line intensities ... 330
1.4.2. Temperature evaluation from continuous radiation 332
1.5. Some plasma properties at high temperatures ... 333
1.5.1. The composition of the plasma ... 333
1.5.2. The dynamical viscosity of a plasma ... 334
1.5.3. The electrical conductivity of a plasma ... 337
1.5.4. The specific heat of a plasma and the thermal conductivity ... 338

§ 2. The production of high temperature by means of the electric arc ... 342
2.1. The arc technique ... 342
2.1.1. The free-burning arc ... 342
2.1.2. The low-voltage high-current arc ... 343
2.1.3. The Beck arc ... 344
2.1.4. The Gerdin diaphragm ... 345
2.1.5. The water-pipe arc ... 346
2.1.6. Recent development of high temperature arc technique ... 349

PAGE

2.2. The scientific use of the high temperatures obtained by arcs 349
2.2.1. The evaluation of transition probabilities 349
2.2.2. The evaluation of collision cross sections 350
2.2.3. The evaluation of plasma properties 351
2.3. The practical use of the high temperatures obtained by arcs 352
2.3.1. The use of arcs in chemical industry 352

§ 3. High temperatures obtained with pulsed discharges and sparks 352
3.1. Methods for observation of pulsed light sources 352
3.2. Temperatures measured in pulsed discharges 353
3.3. The effect of pressure waves in pulsed discharges 354
3.4. The increase of temperature by the pinch effect 355
3.5. Temperature evaluation in sparks 357
3.6. Recent development of spark technique for obtaining high temperatures 358

§ 4. High temperatures obtained with exploding wires 358
4.1. The fundamental processes related to wire explosion 358
4.2. The evaluation of the temperature reached in the two stages of the explosion 360
4.3. Further development of the exploding wire technique for obtaining high temperatures 361

§ 5. High temperatures obtained with plasma jets and shock waves 363
5.1. High temperatures obtained with plasma jets 363
5.1.1. Plasma jets within arcs 363
5.1.2. The temperature of plasma jets ejected from arcs 364
5.1.3. The temperature of plasma jets ejected from arcs with supersonic velocity 365
5.2. High temperatures obtained with plane shock waves 366
5.2.1. Plane shock waves generated by sparks 366
5.2.2. Plane shock waves generated in the shock tube 367
5.3. Further development to obtain higher temperatures with shock waves 369
5.3.1. Shock waves of different geometrical form 369

§ 6. High temperatures obtained with chemical reactions 370
6.1. High temperatures obtained with detonations 370
6.1.1. The three mechanisms of reaction: combustion, explosion, detonation 370
6.1.2. The hydrodynamic theory of the plane detonation wave 371
6.1.3. Spectroscopic measurements of detonation temperatures 372
6.1.4. Temperatures of the detonation front 373
6.1.5. Converging detonations 375
6.2. High temperatures obtained with shock waves from detonations 376
6.2.1. Shock wave generation by detonation 376
6.2.2. The luminosity of shock waves 378
6.2.3. Shock waves from nuclear explosions 379

References 380

Abstract. The article is based mainly upon the study of thermally excited plasmas with temperatures in the range of 10 000–70 000°K. In case of no self-absorption *the measurement* of these temperatures can be effected by observation of lines or of continuous radiation emitted from the plasma. The study of lines implies both line intensities and line shapes.

Following the discussion of the different methods of evaluating temperatures the specific properties of a plasma—viscosity, electrical and heat conductivity—are touched upon.

Sections 2 to 6 deal with the *production* of high temperatures. Arcs, pulsed discharges and sparks, exploding wires and shock waves, and finally chemical reactions are discussed. The temperatures actually obtained, the practical applications, and the limitations of the different approaches are indicated.

The writing of the article was completed early in 1957 but some of the most interesting results in the field of extremely high temperatures obtained during 1957 were added at the proof stage. The methods of temperature measurement presented here have been used frequently by many investigators and have proved to lead to reliable results of known accuracy in the temperature range below 10^5 deg K. It is expected that the same methods with minor alterations could be used in the region of temperatures exceeding 10^8 deg K.

---

§1. Measurement of High Temperatures

1.1. *Introduction*

The production and observation of high-temperature phenomena have always been of particular interest to the human mind. The temperature range considered to be high, however, was different at different times. Disregarding earlier history, the heat of the oxy-hydrogen flame and the heat developed by the electric arc between carbon rods have been known and thought of as very high temperatures during the past century. Indeed, the temperatures thus obtained are the highest which solid matter can stand. It was not until the beginning of the present century that the field of accessible temperatures widened considerably due to a change of interest from heated solid matter to the behaviour of hot gases. Whilst the sublimation of carbon at about 4000°K puts a limit to the temperature which solid matter can stand, gaseous matter may fairly easily be heated to a sustained temperature of 20 000 to 50 000°K, whilst transient temperatures of 10^6 or even 10^7 degrees Kelvin have apparently been obtained recently. This rapid increase in the range of accessible temperatures makes it necessary to concentrate firstly on our present conception of temperature. It will be seen that our ideas of the time required to reach temperature equilibrium have lately changed markedly and have led to an understanding of the very high transient temperatures which have been reported. Secondly, it should be emphasized that almost all measurements of temperatures exceeding 2000°K are based upon optical observation. Therefore we have to discuss the mutual influence of matter and radiation in order to be able to evaluate the temperature from observation either of continuous or of line emission. In the later sections of this article we shall describe methods of obtaining sustained temperatures of 10 000 to 50 000°K and methods of obtaining transient temperatures from 10 000°K up to very much higher temperatures. All these methods refer to laboratory experiments outlined under well-defined conditions. We shall discuss methods and give examples rather than aim at completeness.

1.1.1. *Temperature and temperature equilibrium.*

The temperature of a gas depends on the condition that energy of every particular kind is distributed over all the different particles present in the gas according to a Boltzmann distribution. The energy may be carried by the

different particles as translational, rotational or vibrational energy, as energy of excitation and ionization, or may be energy of radiation in between the particles. The definition of a state of equilibrium implies that the number of excitations of atoms by electrons is equal to the number of collisions of the second kind, that events of ionization by collision are equal to the events of recombination between three colliding particles and that the radiation emitted is equal to the radiation absorbed. In hot gases, instead of such a state of equilibrium, we can only realize some kind of a stationary state determined by the fact that a constant supply of energy balances a number of inevitable energy losses, especially losses by convection, by diffusion, and by radiation.

Most of the different kinds of energy change and transform mutually by collisions. Experiments have shown that reduction of the pressure to a few millimetres of mercury reduces the number of collisions so much that no uniform Boltzmann distribution results. However, at atmospheric or higher pressure a uniform stationary Boltzmann distribution has been observed, provided the energy loss due to radiation is not too large. These experimental results refer to flames outside the reacting zone just as much as to electric arcs. It is most remarkable that the emission from the plasma of an electric arc may be treated as an equilibrium radiation corresponding to the temperature of the plasma in spite of the complicated mechanism of the electric discharge. Experiments to control the thermal behaviour of an electric arc have been carried out carefully at various places, so that we may take this equilibrium for granted (Compton 1923, Mannkopff 1932, Ornstein and Brinkman 1934, Kruithof 1943 a, b, Kruithof and Smit 1944).

1.1.2. *Time required to establish temperature equilibrium.* (See also § 1.1.3.)

The time required to establish thermal equilibrium depends upon the number of collisions between the different particles. In the case of electrical discharges we note that electrons drifting in a field of given direction and colliding with each other will keep the drift velocity of their centre of gravity regardless of the number of collisions. Therefore no increase of the random velocity distribution of the electrons will be observed, i.e. no rise of temperature. Only collisions with heavy gas particles will induce the electrons to change their directions independently. A few collisions with heavy particles will then suffice to produce random velocity distribution and hence an increased electron temperature.

If we denote the time between collisions of electrons with gas atoms by τ we obtain

$$\tau = \frac{1}{v_e(n_0 Q_0 + n_i Q_i)} \qquad \ldots\ldots(1)$$

where v_e is the thermal velocity of the electrons, n_0, n_i the number density of neutral and ionized atoms respectively, Q_0 the Ramsauer cross section of atoms against electrons, and Q_i the ionic cross section (Gvosdover cross section).

Our knowledge of cross sections for electron collisions has recently been considerably extended, as will be seen from § 2.2.2. We may note here that generally the Ramsauer cross sections are considerably larger than gas kinetic cross sections, whilst the cross sections of ions in many cases are larger still. As an example we evaluate the time τ for hydrogen gas at atmospheric pressure and temperature $T = 10\,000°\text{K}$. At this temperature, the number of molecules is negligible whilst

the number of ions is still small. Inserting $Q_0 \simeq 170 \times 10^{-16}$ cm²†, $n_0 \simeq 10^{18}$ cm^{-3}, $v_e \simeq 10^8$ cm sec^{-1} we obtain $\tau \simeq 10^{-12}$ sec. To ensure random velocity distribution we allow for a few collisions and obtain $\tau \simeq 10^{-11}$ sec as the time to establish the electron temperature.

We consider next the interaction between electrons and ions. In a single elastic impact the electrons of mass m give up to the atoms of mass M only a fraction of their surplus energy proportional to $2m/M$. It will take about 10^3 impacts to transfer the surplus energy to the atoms. In the example given we obtain for the time of energy exchange between electrons and hydrogen atoms approximately 10^{-8} sec. The cross section Q_0 for H atom–electron interaction is fairly large. Thus, equilibrium times for other gases may vary between 10^{-5} and 10^{-8} sec.

With increasing temperature the degree of ionization and hence the product $n_i Q_i$ of formula (1) will increase. In the above example the value of Q_i is just about as large as the value of Q_0. Thus for hydrogen and likewise for the alkalis the time of energy exchange does not vary much with temperature. When, however, the Ramsauer cross section Q_0 is small the time of energy exchange decreases considerably with increasing ionization. When highly ionized gas atoms are present the ' degree of ionization ' may be higher than 100% and the time required to attain equilibrium becomes shorter. This will be of importance for transient high temperature phenomena, especially in the case of sparks. For a spark discharge at atmospheric pressure recent research has led to a measured temperature of 40 000°K with an experimentally proven uniform, though time-dependent, equilibrium temperature (see § 3.5). This result—though somewhat surprising—can be explained by taking account of the high degree of ionization in the fairly strong condensed discharge.

Considering the facts more closely in the case of an electric discharge, the electrons are constantly gaining surplus energy from the electric field. Consequently a total transfer of the surplus energy to the atoms will never be realized : the temperature of the electrons remains always somewhat higher than that of the gas. Equilibrium is reached when the energy transfer from the electrons is equal to the energy gain from the field :

$$\tfrac{3}{2} k(T_e - T_g)^2 \frac{m}{M} = eE^2 b_e \tau \qquad \ldots\ldots(2)$$

where

$$b_e = v_e/E = \tau e/m \qquad \ldots\ldots(3)$$

is the mobility of the electrons, and k is the Boltzmann constant. For the relative difference in temperature we obtain from (2) with $\tau = \lambda/v$

$$\frac{\Delta T}{T_e} = \frac{T_e - T_g}{T_e} = \frac{(\lambda e E)^2}{(\tfrac{3}{2} k T_e)^2} \frac{M}{4m}. \qquad \ldots\ldots(4)$$

Considering a free burning arc in air with $T = 6000$°K, $E = 20$ v cm^{-1}, and the mean free path of the electrons $\lambda = 1{\cdot}3 \times 10^{-5}$ cm we obtain $\Delta T = 6$°K or 0·1%.

Observation of line intensities leads primarily to the temperature of excitation and ionization ; at atmospheric or higher pressure both these temperatures are

† The value of $1{\cdot}7 \times 10^{-14}$ cm² is the experimental mean value between those of Drawin (1956) and Maecker, Peters and Schenk (1955).

near the electron temperature. The temperature difference $T_e - T_g$, though small for free burning arcs, increases with the field strength E. Taking $E = 300$ v cm^{-1} as has been found for a constricted water-cooled arc and for high pressure, the difference $T_e - T_g$ becomes noticeable and, in this case, it might be incorrect to combine a measured value of T_g (e.g. from Doppler effect) with measured line intensities depending on T_e.

1.1.3. *Effect of a temperature gradient in a light source.*

Now assume a plasma with a temperature gradient, as is realized in every high temperature light source. It is customary to consider each element of volume separately as having a local thermodynamic equilibrium in itself. Using stigmatic optical analysis temperature gradients have been evaluated. A more detailed analysis however shows that the elements of volume are not quite independent. Particles are transported from one element to another by convection and diffusion. A particle freshly introduced into an element of volume of different temperature will require a certain relaxation time to become adjusted to the new surroundings. In order to estimate this relaxation time the velocities of reaction for the processes of ionization, recombination, dissociation, etc., are required. Although we are far from knowing these data exactly we may conclude that a measured temperature gradient is in error because of relaxation effects, the error being larger, the larger the temperature gradient. In extreme cases it may be worth while to consider relaxation times to obtain a clear understanding of the meaning of the parameters measured and the accuracy obtained.

1.1.4. *Effect of the optical depth of the emitting gas.*

In a hot gas of uniform temperature the emission from some atoms may be reabsorbed by others. Where there is no reabsorption the intensity of the light observed can be related to the emission of a single atom. Some years ago Minkowski (1930) showed that whenever light is strong enough to be photographed reabsorption is noticeable. However, these results apply to temperatures of 1000 to 2000°K with low intensity of the corresponding black-body radiation. Above 10 000°K many light sources are quite transparent even in the central parts of the lines owing to the fact that the intensity of black-body radiation, i.e. the maximum possible intensity of emission corresponding to these temperatures, is very high indeed. Experimental evidence for transparency will be obtained either by variation of the geometrical depth of the emitting layer or by measuring absolute intensities. In the former case an optically thin layer is indicated by variation of the light intensity corresponding exactly to the geometrical dimensions, in the latter case, by an absolute intensity which is only a small fraction of the corresponding black-body radiation. Experimental evidence has shown many light sources operating at temperatures above 10 000°K and at atmospheric pressure to be optically thin. Other light sources may be found to be optically thick in the central parts of some lines but optically thin in other spectral regions.

1.2. *Determination of Temperature*

1.2.1. *The determination of the gas temperature in the case of an optically thick layer.*

In the case of emission from an optically thick layer the temperature can be obtained simply by absolute intensity measurements and application of the Planck

formula. A spectral line which is optically thick at the centre has a profile which is flattened at the top, this part of the line profile being really a section of the black-body radiation curve for which the temperature can be evaluated. Sometimes band heads are especially suited for this kind of temperature evaluation.

When there is a temperature gradient in the line of observation, the frontal layer will absorb some light, causing the well-known phenomenon of self-reversal. This prevents observation of the flattened portion of the line profile. The superposition of an absorption line of *known* profile (the temperature gradient being known) allows the evaluation of the temperature from the maximum intensity of the two halves of a self-reversed line. The profile of the absorption line and the temperature can be obtained from an exposure with the slit of a stigmatic spectrograph perpendicular to the axis of symmetry of the light source. The method cannot be applied to lines involving the ground state and therefore is very appropriate to high temperature research (Bartels 1949 a, 1950 a, b, Göing 1951).

With increased current the continuous background of almost every light source is enhanced relative to the intensities of the lines. Self-reversed lines show up above the continuum as two halves of emission lines with the absorption in between. Further increase of current makes the emission parts almost disappear within the continuum, finally only the absorption line being noticeable. Stigmatic observation of an arc column of cylindrical symmetry, e.g. the column of a pulsed high pressure mercury arc, shows the self-reversal fully developed only in the central line of observation. Parallel to this line decrease of temperature and also of the optical depth is noticed. According to Bartels (1956) the different intensities as a function of distance from the central line of sight give a measure of both the optical thickness and the maximum temperature of light sources with cylindrical symmetry.

When the light source is neither optically thick nor optically thin a reduction to an optically thin layer might be successful. In this case the geometrical depth of the layer should be changed and the observed values of the intensities should be made to fit the curve of growth (Held 1931, Unsöld 1955), from which the reduction to an optically thin layer can be obtained provided the mechanism governing the line-width, the ' damping constant ', is known. This method is especially simple when the line has a multiplet structure. The relative intensities of a multiplet can be made to fit the curve of growth and the reduction to an optically thin layer can be obtained without any further experimental procedure.

1.2.2. *The determination of the gas temperature in the case of an optically thin layer.*

In the case of thermal excitation the intensity distribution over the rotational structure of a single band corresponds to Boltzmann's law. Therefore from intensity measurements over a band the temperature can be derived:

$$I_R(m) = I_P(m+1) \sim \nu^4 \frac{(J+1)^2 - \frac{1}{4}}{J+1} \exp - \left(\frac{h^2}{8\pi^2 \Theta} \frac{1}{kT} J(J+1) \right). \qquad \ldots\ldots (5)$$

In this formula I_R and I_P represent the intensity of an R and P line of a band, m is the corresponding number of the line, J the rotational quantum number, Θ the moment of inertia of the molecule and ν the frequency of the line.

Reliable measurements can only be obtained with the use of a spectrograph of high resolving power and large dispersion. When the dispersion is not large enough the single band lines will overlap to form an apparent continuous intensity distribution. The resulting intensity distribution has been calculated for the CN bands $\lambda = 3883$, 3871, 4216 and 4197 Å (Spier and Smit 1942, Smit-Miessen and Spier 1942, Spier and Smit-Miessen 1942). Using these results it is possible to obtain in a very simple way the temperature even from spectra taken with low dispersion.

It should be noted, however, that in high temperature light sources molecules are dissociated, bands being emitted only from cooler regions of the light source. In each case therefore it will be necessary to make sure that the emission of bands, lines or continuous radiation can really be attributed to the particular element of volume under consideration.

Lines which are not broadened by increased optical depth or by other influences affecting the line width will be found broadened by Doppler's principle. The resulting line profile is represented by the formula

$$\frac{dI}{I} = \frac{1}{\sqrt{\pi}} \exp -\left(\frac{\Delta\lambda}{\Delta\lambda_D}\right)^2 \frac{d\lambda}{\Delta\lambda_D} = \frac{1}{\lambda}\sqrt{\frac{Mc^2}{2\pi RT}} \exp\left\{-\frac{Mc^2}{2RT}\left(\frac{\Delta\lambda}{\lambda}\right)^2\right\} d\lambda \qquad \ldots\ldots(6)$$

where $\Delta\lambda_D/\lambda = \sqrt{(2RT/Mc^2)}$, M being the molecular weight and R the gas constant, whilst $\Delta\lambda_D$ indicates the observed Doppler half-width measured in wavelength units. This formula has been used successfully to evaluate the temperature especially for the lighter elements, the lines of which show relatively large Doppler broadenings. The half-width of a Doppler broadened line depends in a simple way upon temperature :

$$2\Delta\lambda_D = 2\sqrt{(\ln 2)}\,\lambda\sqrt{\frac{2RT}{Mc^2}} = 7{\cdot}16 \times 10^{-7}\lambda\sqrt{\frac{T_g}{M}}. \qquad \ldots\ldots(7)$$

For elements with complex term structure, the incompleteness of inner shells leads to lines which are fairly sharp and thus allows temperature evaluation based on Doppler's principle. As an example, we refer to a temperature evaluation for the free-burning iron arc (Burhorn 1955). Here lines are observed combining energy terms not related to outer electrons. These terms show hardly any perturbations and hence the lines show Doppler broadening. An independent evaluation of the electron temperature T_e results in the same temperature as obtained from Doppler broadening, thus indicating thermal equilibrium.

The two methods mentioned above, together with the method of absorption of either x-rays (von Engel and Steenbeck 1931, 1933, Kenty 1941, Koch 1949) or α-rays (von Engel and Steenbeck 1931, 1933) and finally the method of measuring gas densities either by interferometric methods (Schmitz 1949), by Schlieren methods (Sperling 1950), by measurements of the velocity of sound (Poritsky and Suits 1935, Suits 1935) or by measurements of radioactivity (Ecker 1951), are the basic methods of evaluating T_g, the true gas temperature.

1.2.3. *The determination of the excitation and ionization temperature.*

In the case of a hot plasma we remember that a certain number of electrons will be freed from atoms. The dependence of the degree of ionization upon

temperature can be obtained from Saha's equation (Eggert 1919, Saha 1920, 1921) the law of mass action for the process of ionization :

$$\frac{n_+ n_e}{n_0} = f(T_e). \qquad \ldots\ldots(8)$$

Furthermore, the total free charge in case of thermal excitation always amounts to zero. Thus

$$n_e = n_+. \qquad \ldots\ldots(8a)$$

Finally the concentration of electrons, ions and neutral atoms add to give the total number density N of the gas particles :

$$n_e + n_+ + n_0 = N. \qquad \ldots\ldots(8b)$$

Thus for a thermally excited monatomic gas of known number density (pressure) the three preceding equations, based on equilibrium conditions, are available for evaluation of the four unknowns n_e, n_+, n_0, the number of electrons, ions and neutral gas atoms per cm^3 respectively and the temperature T_e.

Any measurement connected with any one of the four unknowns will provide a fourth equation and thus allow a solution of the set of equations. We obtain the temperature as one of the four quantities of prime importance to high temperature research.

There is hardly any difference when the plasma is composed of more than one constituent. Suppose the relative abundances of the plasma constituents are known, then again just one measurement will permit a solution of the set of available equations. If the relative abundances are not known, as many measurements are required as there are constituents of the gas.

As an example, we take a hot plasma composed of air and gaseous carbon such as is found in the column of a free burning carbon arc. Here, the relative abundance of oxygen to nitrogen is known, but the abundance of carbon is not known. The equations available are three Saha equations for the three constituents of the gas, the charge equilibrium, Dalton's law, and the relative abundance of nitrogen to oxygen, six equations in all :

$$\frac{n_+ n_e}{n_0} = f(T_e) \text{ for N, O, C} \qquad \ldots\ldots(9a)$$

$$n_{N^+} + n_{O^+} + n_{C^+} = n_e \qquad \ldots\ldots(9b)$$

$$n_{N^+} + n_N + n_{O^+} + n_O + n_{C^+} + n_C + n_e = N \qquad \ldots\ldots(9c)$$

$$n_{N^+} + n_N = 3{\cdot}72\,(n_{O^+} + n_O). \qquad \ldots\ldots(9d)$$

The eight unknowns are n_+ and n_0 for three gases, n_e and T. The evaluation of eight unknowns necessitates two measurements. They may be of quite different nature. If there is line emission from the plasma out of an optically thin layer the intensities of lines of known transition probabilities might be measured. In this case the completing equations are of the form

$$I = A_{0,m}\, n_m\, h\nu = A_{0,m}\frac{g_m}{u}\, n_0\, h\nu \exp\left(-\frac{E_m}{kT}\right) \qquad \ldots\ldots(10)$$

where I is the radiation emitted by 1 cm^3 in ergs per second in all directions, n_m and n_0 are the numbers of atoms per cm^3 in the excited state m and of all

neutral atoms, and $A_{0,m}$ is the transition probability for the transition m (upper) $\rightarrow$ 0 (ground state). In the case of thermal equilibrium the number of atoms in the state m is given by the Boltzmann law and the second equality sign of equation (10) is valid.

Equation (10) relates the measured line intensity I to the temperature T_e. Whenever $T_g = T_e$ the set of equations (9*a–d*) can be solved. Instead of measuring an absolute intensity the intensity ratio of two lines of known transition probabilities can be measured. Equally well the absolute intensity of the continuous radiation or the intensity ratio of the continuous radiation at different wavelengths can be measured (see §§ 1.3.6 and 1.4.2). In some cases quite different measurements are preferable, e.g. the current density of an electric discharge or, for shock waves, the pressure ratio across the shock. We will leave these more specialized methods of temperature evaluation to the discussion of the corresponding phenomena.

However, the measurements may just as well concern some other unknown of the set of equations (9*a–d*). Besides T_e, the electron concentration n_e is of major importance; n_e is often related to line broadening and may be obtained by measurement of line width, line profile or line shift as will be seen from the following section.

1.3. *The Use of Line Broadening for the Determination of the Number Density n_e of the Electrons*

1.3.1. *Line broadening by impact.*

In the theory of impact (Lorentz 1905, Lindholm 1941, 1942, 1945, Burkhardt 1940) the emitting atom is supposed to radiate the unperturbed frequency except for the very short moment of collision. These collisions cause a phase distortion of the frequency to be emitted which, as Fourier analysis shows, is equivalent to a broadening of the line.

We consider first just one perturbing particle at a distance r from the radiating atom. The perturbation causes a line (term) shift

$$\Delta\omega = \frac{2\pi C}{r^n} \qquad \ldots\ldots(11)$$

where ω is the angular frequency emitted, C is a constant and r is the distance of the perturbing particle; in particular $r_{\min} = \rho$ is the collision parameter.

In this formula n depends on the nature of the perturbation: $n = 2$ represents the perturbation by a charged particle of an atom whose levels show a linear Stark effect; $n = 3$ represents the mutual interaction of identical neutral particles, also the action of a charged particle and atoms having quadrupole moments; $n = 4$ represents the perturbation by a charged particle of atoms showing a quadratic Stark effect; $n = 6$ represents the interaction of neutral particles by van der Waals forces.

The perturbation integrated over different values of ρ results in any case in a line of dispersion profile

$$I(\omega) = \frac{1}{(\omega - \omega_0 + \Delta)^2 + (\delta/2)^2}. \qquad \ldots\ldots(12)$$

The term Δ indicates the line shift whilst δ is the half-width of the line.
For $n = 4$ we obtain

$$\delta = 38{\cdot}8\, C^{2/3} v^{1/3} n_e \qquad \ldots\ldots(13)$$

and

$$\Delta = 33{\cdot}4\, C^{2/3} v^{1/3} n_e\,; \qquad \ldots\ldots(14)$$

the ratio δ/Δ, i.e. the ratio of half-width to displacement, gives the value 1·16 (see Unsöld 1955).
For $n = 6$ we obtain

$$\delta = 17{\cdot}0\, C^{2/5} v^{3/5} n_e \qquad \ldots\ldots(15)$$

and

$$\Delta = 6{\cdot}76\, C^{2/5} v^{3/5} n_e\,; \qquad \ldots\ldots(16)$$

the ratio δ/Δ amounts in this case to 2·8 (see Unsöld 1955). It is possible to decide quite easily from experimental measurements of half-width and shift of a line between cases $n = 4$ and $n = 6$, i.e. between the influence of quadratic Stark effect and van der Waals forces, which otherwise might be difficult to discriminate. If quadratic Stark effect and van der Waals forces act simultaneously, the displacements caused by each action separately will add so that $\delta_{res} = \delta_4 + \delta_6$ and $\Delta_{res} = \Delta_4 + \Delta_6$. From this we obtain

$$\Delta_{res} = \frac{\delta_4}{1{\cdot}16} + \frac{\delta_6}{2{\cdot}8}. \qquad \ldots\ldots(17)$$

This means that the resultant and experimentally observed line displacements can be split into a sum of two terms, corresponding to the half-widths caused by quadratic Stark effect and by van der Waals forces. Having separated the two effects we obtain the value of n_e from the quadratic Stark effect using formula (13) or (14). Although this method of evaluating n_e is not very accurate, Mandelstam (1956 b) has obtained the same value of n_e from very different plasma lines using interferometry to measure both half-widths and displacements.

It should be mentioned, however, that when $n = 4$ the ratio of δ/Δ found experimentally is often different from the theoretical value 1·16. This probably indicates that the interaction in these cases was not quite quadratic.

1.3.2. *Statistical line broadening.*

Whilst the perturbations discussed above were based on the assumption that the emitting atom was perturbed by impact only, in the statistical theory the emitting atom is supposed to radiate always in a perturbing field of surrounding charges.

(*a*) *Perturbation by one particle.* The perturbation by one particle can be calculated from the probability that an (always) perturbing particle is found at a distance r. For the purpose of calculation the emitting atom is thought of as surrounded by a sphere and spherical shells, which are analysed with respect to the probability of finding a perturbing particle inside such a shell. This probability is related to the intensity distribution, i.e. the profile of the line

$$I(\omega)d\omega = \frac{3}{n}\left(\frac{\Delta\omega_0}{\Delta\omega}\right)^{(3/n)+1} \exp\left\{-\left(\frac{\Delta\omega_0}{\Delta\omega}\right)^{3/n}\right\}\frac{d\omega}{\Delta\omega_0} \qquad \ldots\ldots(18)$$

where $\Delta\omega_0 = 2\pi C/r_0^n = 2\pi C(\frac{4}{3}\pi N)^{n/3}$ is the displacement caused by a particle in the mean distance r_0 defined by $\frac{4}{3}\pi r_0^3 N = 1$ and as before $\Delta\omega = 2\pi C/r^n$.

The ' single particle approach ' (equation (18)) is of interest only for large $\Delta\omega$ because that is the case in which one perturbing particle is close to the atom under consideration. Consequently the exponential in formula (18) is close to unity and may be neglected.

(*b*) *Perturbation by several particles.* Next we consider the perturbation imposed on the emitting atom not only by one but by several perturbing particles at the same time. For $n = 2$ and $n = 4$, i.e. perturbation by *charged* particles, the electrical fields from several perturbing particles add vectorially. Again a probability function W of the strength of the resulting microfield can be calculated, as has been done by Holtsmark (1919 a, b, 1924), Verwey (1936), Kuhn (1934), Margenau (1932) and Schmaljohann (1936). From the probability function the intensity distribution of the lines is derived.

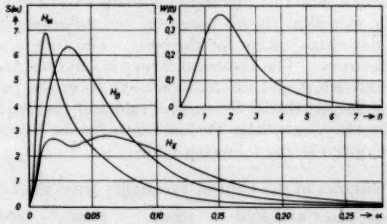


Figure 1. The profiles of the Balmer lines H_α, H_β and H_γ according to the statistical theory and the probability function $W(\beta)$. $\alpha = \Delta\lambda/F_0$, $\beta = F/F_0$.

For convenience, the field strength F has been expressed in terms of a ' normal ' field strength

$$F_0 = 2{\cdot}61eN^{2/3}, \qquad \ldots\ldots(19)$$

where N is the number of perturbing charges per cm³. The method is particularly appropriate to lines showing linear Stark effect, e.g. the Balmer lines, which show a particularly broad line profile. To obtain the number of perturbing particles N, we draw a number of line profiles as shown in figure 1, which have been calculated for different values of F_0. We choose the ' theoretical ' curve which fits best with the experimental profile and obtain N from the corresponding value of F_0.

It may be noted that this method of obtaining N the number of perturbing charges per cm³ is simple and fairly accurate, because the linear Stark effect causes large splittings and correspondingly broad half-widths of Balmer lines. An additional advantage of this method of evaluating N lies in the fact that hydrogen lines are present in almost every flame or electrical discharge, in the latter case diffusing in from the interior of the electrodes or from the walls. Equally well hydrogen can be introduced easily with moist gases. The Balmer lines then act as a kind of ' N-indicator '.

The disadvantage of the method is the fact that the meaning of ' the number of perturbing particles ' is by no means clear. To elucidate this we consider the assumptions of Holtsmark's theory. The perturbed radiating atom emits a

line shifted about $\Delta\omega$ with a half-width about $1/\tau$ (τ is the time of practically constant perturbation $\tau \simeq \rho/v$; v is the velocity of the perturbing particle). A number of shifted and broadened lines will add according to their frequency of occurrence, as long as

$$1/\tau \ll \Delta\omega \qquad \ldots\ldots(20)$$

and the statistical line profile results (Lenz 1933).

For $1/\tau$ greater than $\Delta\omega$, i.e. the displacement during the collision smaller than the line width, the impact theory can be applied. Expressing τ by ρ/v and $\Delta\omega$ by C/ρ^2, i.e. for linear Stark effect, we obtain for equation (20)

$$C/\rho v \gg 1. \qquad \ldots\ldots(20a)$$

This indicates that the statistical theory holds for small values of ρ and also of v. The slowly moving ions fulfil this condition for a larger value of ρ, i.e. for a smaller value of $\Delta\omega$. The statistical theory based on the influence of ions allows the calculation of the line shape except for the core. The core is mainly influenced by impact of the electrons. The statistical theory is valid for $\Delta\omega > v^3/2\pi^3 C$ in the case of a linear Stark effect and for $\Delta\omega > (8v^4/\pi^5 C)^{1/3}$ in the case of a quadratic Stark effect, and the impact theory for smaller values of the frequency shift. The distance of the critical frequency from the line centre for the Balmer lines according to Unsöld (1955) is given in the following table:

Table 1. The Distance of the Critical Frequency from the Line Centre in Å

Temperature (°K)		25000	10000	5000	3000
H_α	electrons	580	230	110	70
	H ions	0·63	0·25	0·12	0·08
H_β	electrons	120	48	24	14
	H ions	0·13	0·05	0·03	0·02
H_γ	electrons	48	19	9	6
	H ions	0·05	0·02	0·01	0·006
H_δ	electrons	32	13	6	4
	H ions	0·03	0·01	0·007	0·004

Nearer to the centre of the line only the ions give statistical broadening, whilst the action of the electrons *though noticeable* cannot be calculated by statistical methods. Further out the tails of the lines are influenced by ions and electrons. Whilst the participation of the electrons for large $\Delta\omega$, i.e. in the tails of the lines, seems fairly certain, there is a further limitation to the action of the electrons, which is more difficult to calculate, i.e. the limit of adiabatic line shift. If the perturbing particle is too fast the Stark splitting of the levels does not quite follow the rapidly changing outer field, or properly speaking the eigenfunctions of different Stark levels intermix, thus setting a limit to the theories given above.

1.3.3. *Experimental proof of Holtsmark's theory.*

The safest procedure is an experimental proof, comparing experimentally obtained profiles of Balmer lines with profiles obtained by calculation in the statistical way. But this is not enough. We require an independent evaluation of n_e for comparison with the value obtained for N. A careful study of this was made by Jürgens (1952). He found the profiles of the Balmer lines to correspond to

curves calculated from statistical methods and in addition the value of N to be equal to n_i, the number of ions per cm^3, i.e. no participation of the electrons. However, his results called for reconfirmation as they were obtained from an arc viewed ' side on ', so that a reduction to a layer of uniform temperature had to be made. The reduction seems to be trustworthy, as long as the light source has a steady position with respect to the observer. For the electric arc used by Jürgens this condition was not fulfilled as will be shown later (§ 2.1.5). Besides, the photographic method used by Jürgens could not cover a very large intensity scale without additional uncertainty. Therefore it was decided to repeat Jürgens' experiment using photoelectric equipment and making ' end-on ' observations on

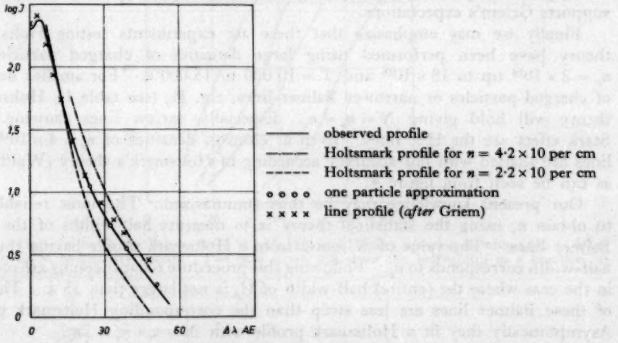


Figure 2. The line profile of H_β.

a layer more uniform with respect to temperature. Bogen (1957, see also Lochte-Holtgreven 1956) in performing these experiments and using values of n_e ranging from 2×10^{15} to 18×10^{15} has never obtained a perfect fit between the observed Balmer profiles and the calculated Holtsmark profiles, the latter being definitely steeper (see figure 2). The agreement, however, is fairly good for the tails of the lines, improving steadily with distance from the centre. The value of N taken from the Holtsmark profile fitting asymptotically to the tails corresponds *not* to the value of n_e found by solving the set of equations (9*a–d*), an absolute line intensity being measured, but to twice the value of n_e, indicating the simultaneous activity of the electrons!

Two attempts have been made recently to include the action of the electrons in the theory. Griem (1954) has considered the levels split by the Stark effect of the ions. These single levels will be broadened by impact from electrons. Similar calculations were made by Kolb (1956). But while Griem assumes the dipole axis of the perturbed atoms to face the passing electrons at all times, Kolb assumes the atoms to be aligned by the ionic field, the electrons acting only in strengthening or weakening this field. Both theories derive a smoother slope for the tails of the lines and in addition a somewhat larger line-width. The

theories differ in that Kolb expects the influence of the electrons to be noticeable even at small electron densities, while according to Griem the line profile is not influenced by the electrons for small electron densities, this influence becoming noticeable at higher densities only.

It is quite reasonable to interpret our experimental results (figure 2) as a deviation from the Holtsmark profile, the line tails being less steep. In this respect the two new theories are fully supported.

A further investigation was begun on the half-widths of the different Balmer lines as a function of temperature. It was found that for low temperatures the measured half-widths show a different dependence on temperature from that for higher temperatures (figure 3) (Lochte-Holtgreven 1957). This result supports Griem's expectations.

Finally we may emphasize that these arc experiments testing Holtsmark's theory have been performed using large densities of charged particles, i.e. $n_e = 2 \times 10^{16}$ up to 18×10^{16} and $T = 10\,000$ to $15\,000°\text{K}$. For smaller densities of charged particles or narrower Balmer lines, e.g. H_α (see table 1), Holtsmark's theory will hold giving $N = n_i = n_e$. Especially narrow lines showing linear Stark effect are the He^+ lines. Even at electron densities of $n_e = 4 \times 10^{16}$ these lines are shaped with fair accuracy according to Holtsmark's theory (Wulff 1958) as can be seen from figure 4.

Our present knowledge may be thus summarized. The most reliable way to obtain n_e using the statistical theory is to measure half-widths of the lower Balmer lines. The value of N found from a Holtsmark profile having the same half-width corresponds to n_e. Following this procedure reliable results are obtained in the case where the (entire) half-width of H_β is not larger than 35 Å. The tails of these Balmer lines are less steep than the corresponding Holtsmark profile. Asymptotically they fit a Holtsmark profile with $N = n_i + n_e = 2n_e$.

The bad fit of Holtsmark's theory is not very satisfactory. Unfortunately there is little hope of getting better agreement with theory using quantum mechanical calculations. Margenau (1957) has tried to calculate the broadening influence of the electrons on the different Stark sub-levels by impact theory treating the electrons as waves. Unfortunately quantum mechanics yields half-widths which are definitely in excess of the experimental values.

1.3.4. *The evaluation of n_e from the apparent series limit.*

The broadening of the Balmer lines by Stark effect increases with the quantum number m of the line observed. On the other hand the distance of neighbouring lines decreases. As Inglis and Teller (1939) have shown, the broadened lines near the series limit form an apparent continuum, thus shifting the true series limit apparently to longer wavelengths. Quantitative formulation of this effect gives for the Balmer line of maximum quantum number to be observed at the border of the apparent continuum the relation

$$\log N = 23{\cdot}16 - 7{\cdot}5 \log m_{\max} \qquad \ldots\ldots(21)$$

where $N = n_e + n_i = 2n_e$ for $T < 10^5/m_{\max}$ and $N = n_i = n_e$ for $T > 10^5/m_{\max}$. This formula has been derived as an approximation for astrophysical purposes, where the number of lines observed is large. A fairly detailed study of the number

of Balmer lines observed in a radiating plasma has been made, determining in each case the electron density by independent methods (Nissen 1953). It has been shown, that even when the electron densities are large, and accordingly only

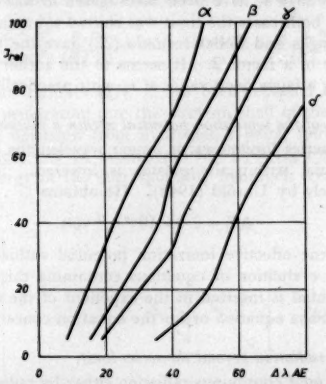


Figure 3. Half-widths of the Balmer lines H_α, H_β, H_γ and H_δ as a function of intensities. (*After* von Ende, preliminary results. The final curve will show the half-widths as a function of the electron density n_e.)

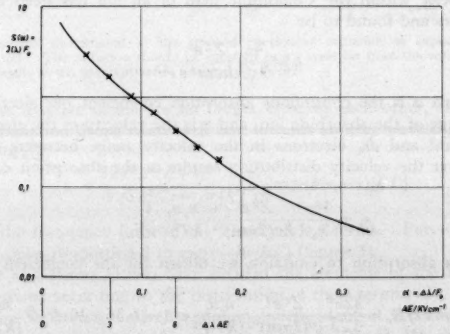


Figure 4. The profile of the He II line $\lambda = 4686$ Å. Full curve: profile according to the statistical theory. Experimental values indicated by crosses.

two or three Balmer lines are left, the Inglis–Teller formula may still be successfully applied. A slight modification of the constant term leads to almost perfect agreement with experiment:

$$\log N = 23{\cdot}46 - 7{\cdot}5 \log m_{\max}. \qquad \ldots\ldots(22)$$

We remember now that the higher terms of any element are hydrogen-like. One expects therefore similar effects near the series limits of other elements. Up to the present the series limit of helium, situated at $\lambda = 2600$ Å and the series limit of magnesium at $\lambda = 2514$ Å, have been investigated in this respect (Wulff 1958, Bötticher 1957). In both cases the limit was shifted towards longer wavelengths. Application of the Inglis and Teller formula (22) gave the true electron density with a possible error of a factor 2. It seems to the author that this method is suitable for obtaining a fairly good value of n_e with practically no effort.

1.3.5. *The diminution of the ionization potential within a plasma.*

The shift of the series limit towards longer wavelengths is an indication that the ionization potential within the plasma is lowered. This effect was first estimated quantitatively by Unsöld (1948). He obtains

$$\Delta E_i = 7{\cdot}0 \times 10^{-7}\, n_e^{1/3} \text{ volt.} \qquad \ldots\ldots(23)$$

This diminution of the effective ionization potential within the plasma is very important indeed for evaluation of equations containing this quantity, especially if the ionization potential is inserted in the exponent of the exponential function as is the case for the Saha equation or for the equation concerning line intensities.

1.3.6. *The continuous radiation beyond the series limit.*

A hot plasma will emit continuous radiation either by radiation of recombining electrons or by radiation of free electrons, i.e. by free–bound or free–free radiation of electrons. Turning first to absorption, the free–free absorption of an electron of given velocity within the Coulombic field of an ion has been calculated by H. A. Kramers and found to be

$$a = \frac{4\pi}{3\sqrt{3}} \frac{Z^2 e^6}{hcm^2 v} \frac{1}{\nu^3}. \qquad \ldots\ldots(24)$$

In this formula a is the continuous absorption coefficient per electron and ion, Ze is the charge of the absorbing ion, and v is the velocity of the electron. With n_i ions per cm^3 and dn_e electrons in the velocity range between v and $v+dv$ integration over the velocity distribution results in the absorption coefficient per centimetre :

$$\kappa_\nu = \frac{16\pi^2}{3\sqrt{3}} \frac{Z^2 e^6}{hc(2\pi m)^{3/2}} \frac{n_e n_i}{(kT)^{1/2}} \frac{1}{\nu^3}. \qquad \ldots\ldots(25)$$

Turning from absorption to emission we obtain for the coefficient of free–free emission

$$\epsilon_\nu = \kappa_\nu(1-e^{u})\, B_\nu = \frac{32\pi^2}{3\sqrt{3}} \frac{Z^2 e^6}{c^3(2\pi m)^{3/2}} \frac{n_i n_e}{(kT)^{1/2}} e^{-u} = 6{\cdot}36 \times 10^{-47} Z^2 \frac{n_i n_e}{(kT)^{1/2}} e^{-u} \qquad \ldots\ldots(26)$$

with $u = h\nu/kT$. The intensity of the free–free emission as given in equation (26) increases towards longer wavelengths. We turn now to the free–bound emission. In the case of broadened lines the continuous intensity distribution beyond the series limit extends steadily to longer wavelengths. Therefore the integration can be extended into the region of discrete terms as far as these are situated

close to each other. Out of recombination from each term a continuous radiation extends with increasing intensity towards shorter wavelengths. This radiation is superposed by the free–free radiation resulting in a total continuous radiation of equal intensity within each interval of frequency (Unsöld 1938):

$$\epsilon_\nu = \frac{32\pi^2}{3\sqrt{3}}\frac{e^6(Z+s)^2}{c^3(2\pi m)^{3/2}}\frac{n_e n_1}{(kT)^{1/2}} = 6{\cdot}36\times10^{-47}(Z+s)^2\frac{n_e n_1}{(kT)^{1/2}}. \qquad \ldots\ldots(27)$$

As will be seen from formula (27), Z has been replaced by $Z+s$, because the electrons are partly penetrating into the electron shell of the ions. The quantity $Z+s$ represents the deviation from hydrogen, s being fairly small. It may be noted that $(Z+s)^2$ may not be taken from the Rydberg formula $m^2(E_H-E_m)/E_H$ where E_H is the energy of the corresponding hydrogen term. The actual value is decidedly smaller.

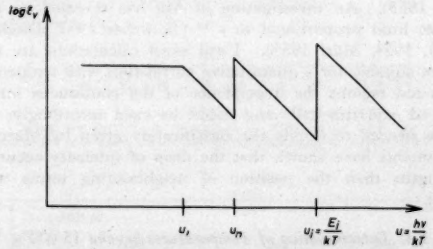


Figure 5. Intensity distribution of the emitted continuous radiation as expected from theory (schematic). The radiation should be constant in a ν scale far from the series limit and drop exponentially at the different series limits. $u = h\nu/kT$.

When the discrete terms are no longer close to each other, the intensity of continuous radiation drops from each line towards shorter wavelengths according to $\exp(-h\nu/kT)$:

$$\epsilon_\nu = \frac{32\pi^2}{3\sqrt{3}}\frac{e^6(Z+s)^2}{c^3(2\pi m)^{3/2}}\frac{n_e n_1}{(kT)^{1/2}}\exp[h(\nu_1-\nu)/kT] \qquad \ldots\ldots(28)$$

where ν_1 is the frequency limit of the closely lying terms. For $\nu > \nu_1$ we obtain an intensity drop proportional to $\exp(-h\nu/kT)$ (figure 5).

Maecker and Peters (1954) have divided the elements of the Periodic Table into three groups according to the distribution of their terms. In the first group the terms are separated, and an emission independent of ν is only expected in the infra-red spectral region. To this group will belong H, He, groups I and II and part of group III of the Periodic Table. The second group has closely lying terms except for the far ultra-violet. This group includes the rare gases with the exception of helium, groups VII, VI, V and part of group IV of the Periodic Table. Finally in the third group we find the elements with highest multiplicity. Here the terms follow closely right up to the series limit of the main series, i.e. over all the visible and near ultra-violet spectral regions.

A number of continuous spectra has been shown to correspond to the above theory. The free-burning arc between carbon rods shows a continuous radiation of equal intensity, when the intensity is represented as a function of the frequency (ter Horst and Rutgers 1953). This radiation extends between $\lambda = 6000$ Å and $\lambda = 4000$ Å and has been attributed to N and O atoms. If a frequency is chosen for which formula (27) is valid, then from the measured emission ϵ_ν the temperature may be evaluated. This has been done by Burhorn (1956); his results show that the temperature evaluated according to (27) fits well with the values of temperatures evaluated from lines. We shall draw attention to the fact that for numerous elements the drop of intensity below the series limit is not proportional to $\exp(-h\nu/kT)$ but follows a different law. From quantum mechanical calculation helium shows an intensity drop below the series limit according to $1/\nu^2$ instead of $1/\nu^3$ (in absorption). Very different slopes have been observed for the continua following the main series of the metals (Ditchburn, Jutsum and Marr 1953, 1954, see also Marr 1955). An investigation of Mg has revealed an intensity drop below the series limit proportional to ν^{-14} (Bötticher 1957, Ditchburn, Jutsum and Marr 1953, 1954, Marr 1955). Until exact calculations are available these continua are not suitable for a quantitative correlation with temperature. In the visible or infra-red regions the dependence of the continuous intensity upon ν should be proved experimentally and might be used accordingly. Furthermore experiments are needed to decide the classification given by Maecker. At least for iron, experiments have shown that the drop of intensity occurs at markedly longer wavelengths than the position of neighbouring terms would indicate (Burhorn 1955).

1.4. *Determination of Temperatures beyond* 15 000°K

1.4.1. *Temperature evaluation from line intensities.*

We return to the intensity of a thermally excited spectral line

$$I = A_{0m}\, n_m\, h\nu = A_{0m}\, n_e\, h\nu \exp\left(-\frac{E_m}{kT}\right). \qquad \ldots\ldots(10)$$

For higher temperatures the exponential function approaches unity whilst the number of atoms in the ground state n_e decreases rapidly for two reasons, firstly, a portion of the gas is already ionized and, secondly, the gas density is decreasing (the pressure being kept constant). The exponential function and the number of neutral atoms counteract with increasing temperature so that the intensity of every line has a maximum at a certain temperature. This maximum can easily be calculated. The result of these calculations is represented in figure 6 for a few lines.

From figure 6 we can draw a number of conclusions. If a plasma is maintained at a high temperature with a temperature gradient perpendicular to the central line of observation, the intensity of spectral lines will change parallel to this line according to the temperature. The maximum intensity of arc lines is found in the cooler parts, and that of the spark lines in the hotter parts. Thus some parts of the plasma indicate the corresponding temperature directly by an intensity maximum of a certain spectral line. From these fixed points a temperature profile can be derived (Larenz and Bartels 1950, 1951). The method corresponds to the

well-known classification of stars into a spectral sequence. According to Fowler (1925, 1936) and Fowler and Milne (1923, 1924) we obtain lines of maximum intensity for the following temperatures :

Si I	5500°K	N II	19000°K	N III	28000°K
H	10000	Si III	19000	He II	35000
Si II	11000	O II	20000	O III	26000
C II	16000	C III	25000	C IV	40000
He I	17000	Si IV	27000		

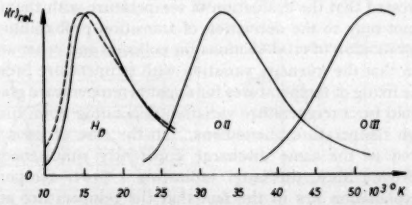


Figure 6. Intensities of the lines H_β, O II $\lambda = 4651$ Å and O III $\lambda = 3447$ Å as a function of the temperature. When the line H_β is broadened by Stark effect from the fluctuating microfields, the broadened line has a somewhat shifted maximum as indicated by the dotted representation.

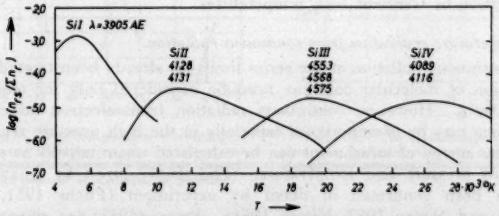


Figure 7. Intensities of some Si lines as a function of the temperature.

In figure 7 we reproduce one of the well-known diagrams from astrophysical textbooks (Payne 1925, Fowler and Milne 1923, 1924). From figures 6 and 7 it will become clear that the accuracy of temperature measurement obtained from the maximum of line intensity is poor. As can be seen from the curves the maximum extends over a region of a few thousand degrees Kelvin giving an uncertainty of 20% or even more in the evaluation of temperature. By far the better method is evaluation of the temperature from the rising slope of the curves of figures 6 and 7. The relative uncertainty in temperature is smaller than the uncertainty in intensity by a factor

$$\frac{\Delta I}{I} = \frac{h\nu}{kT}\frac{\Delta T}{T}. \qquad \ldots\ldots(29)$$

For 10 000°K, $kT \simeq 1$ and hence the factor $h\nu/kT$ is about as large as the excitation energy of the spectral line in ev. From this we conclude that as a 'thermometer'-line we should use a line with a steep rise in the temperature range under consideration, i.e. a different line for every temperature. These lines can be calculated exactly only when the transition probabilities are known; this is the case for the Balmer lines and, for higher temperatures, for the He^+ lines. 'Secondary standards' might be derived from these 'primary standards' by experiment. Experience has shown that fixing of secondary standards is not easy and has to be done repeatedly at different temperatures and for different light sources.

It may be stressed that the evaluation of temperature with the utmost precision is fundamental not only to the derivation of transition probabilities of other lines but also for the evaluation of cross sections for collision and other atomic constants.

We have seen that the intensity variation with temperature indicated in figures 6 and 7 allows the fixing of temperatures for a spatial temperature gradient. Similar considerations hold for a temperature variation depending upon time as is the case for transient high temperature phenomena. In the case of poor time resolution we have observed in the same discharge apparently simultaneous emission of helium and hydrogen lines, the latter indicating a lower temperature than the former. The explanation lies in the fact that the temperature at the beginning of the observation was high, making the helium lines strongest, whilst a little later after a certain drop in temperature the hydrogen lines reached maximum intensity, lines of both elements being recorded on the same photographic plate. This phenomenon may also explain some other inconsistencies which have arisen in the observation of transient high temperatures.

1.4.2. *Temperature evaluation from continuous radiation.*

The continuous radiation at the series limit has already been treated (§ 1.3.6). Superposition of molecular continua need be considered only for temperatures below 15 000°K. However, continuous radiation from electron attachment to neutral atoms may be of importance especially in the high pressure region. For H atoms the energy of attachment can be calculated quantitatively as a function of (electron) pressure and temperature. The theory given by Chandrasekhar (1945) has been confirmed in detail by experiment (Fuchs 1951, Lochte-Holtgreven and Nissen 1952, Nissen 1954). Peters (1953) has shown that for 12 000°K and 150 atm the intensity of the H^- radiation is about 17 times the intensity of both recombination and free–free continua. We conclude from these results that there will be a continuous radiation due to the formation of other negative ions as well, but unfortunately these radiations cannot be calculated and have not yet been measured either. They might, however, play a dominant part in high pressure work.

With high pressure, on the other hand, the re-absorption of radiation is increased. The optical thickness for continuous radiation can be evaluated by the well-known method of increased geometrical size or by comparison with additional emitted lines. Finally for light sources of cylindrical symmetry Bartels' method of stigmatic analysis may be applied (see § 1.2.1).

An estimate of the apparent brightness and temperature of a light source is often obtained by comparison with the radiation emitted from the sun for the

same interval of frequencies. But without further investigation of the optical thickness very doubtful temperature values are obtained.

1.5. *Some Plasma Properties at High Temperatures*

1.5.1. *The composition of the plasma.*

From ionization equilibrium the composition of the plasma can be calculated using Saha's formula (see formula (8))

$$\frac{n_+ n_e}{n_0} = \frac{U_+}{U_0} 2 \frac{(2\pi mkT)^{3/2}}{h^3} \exp(-E_i/kT) \qquad \ldots\ldots(30)$$

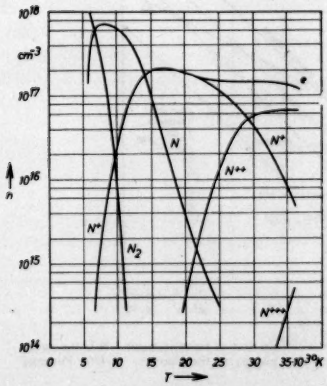


Figure 8. The composition of a nitrogen plasma as a function of the temperature. The number n indicates the number density of particles per cm³ at atmospheric pressure.

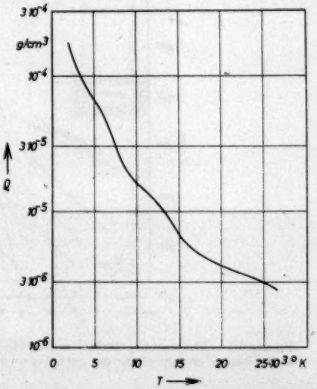


Figure 9. The mass density of a nitrogen plasma as a function of the temperature at atmospheric pressure.

U_+ and U being the partition function (sum of states) for the ion and for the neutral atom. Inserting the values for nitrogen, i.e. for the partition function U_0 the values calculated by Fast (1947), for the energy of ionization the value of 14·54 ev and for dissociation $E_d = 9{\cdot}764$ ev after Gaydon (Gaydon and Penney 1945), we obtain the composition of a nitrogen plasma as shown in figure 8. From the composition the density can be obtained (figure 9). It may be noted that the density does not decrease in a regular way due to ionization effects. As can be seen from the figures the particle densities at high temperatures are low. Even at high pressures the densities are sufficiently low to allow calculations according to the kinetic theory of ideal gases. This becomes clear when the Gvosdover cross section, which represents the sphere of action for charged bodies (see § 1.5.2), is inserted for the 'size of a particle' in the kinetic theory. The critical value is the ratio of the thermal energy to the potential energy of charges with Coulomb interaction. When this ratio is large we are dealing with an ideal gas, whilst a

small value of the ratio indicates a gas with ' plasma interaction '. So far experiments with arcs have not approached the conditions of plasma interaction, as can be seen from figure 10.

1.5.2. *The dynamical viscosity of a plasma.*

With increasing temperature the plasma will show a number of properties which differ decidedly from those of a conventional gas. We consider first the dynamical viscosity η. Following the kinetic theory of gases we obtain

$$\eta = \tfrac{1}{3} nm\lambda\bar{c}. \qquad \ldots\ldots(31)$$

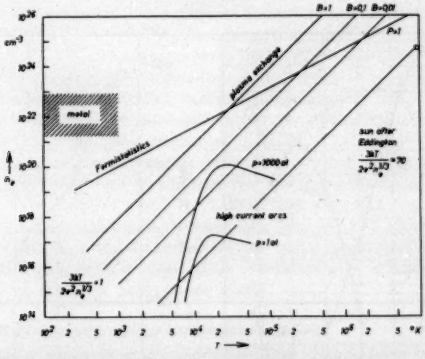


Figure 10. The number density of free electrons as a function of temperature. B is the ratio of the kinetic energy of the gas particles to the potential energy of the charges. (*After* Peters.)

In this formula $\bar{c}$ is the mean velocity of the atoms or molecules, λ the mean free path, m the mass and n the number density of the particles. With $\lambda = 1/nQ$ where Q is the ' collision cross section ' formula (31) becomes

$$\eta = m\bar{c}/3Q. \qquad \ldots\ldots(32)$$

The factor n has disappeared in formula (32) indicating that the viscosity is independent of pressure. Instead, η varies inversely as the collision cross section Q. Some gas kinetic values of Q related to 273°K are: $Q_{He} = 38 \times 10^{-16}$ cm², $Q_{H_2} = 6 \times 10^{-16}$ cm², $Q_{Ar} = 10 \times 10^{-16}$ cm², $Q_{air} = 10 \times 10^{-16}$ cm². It is important to realize that the values of Q change considerably with temperature. In the kinetic theory this change is due to a higher stability of the path of a particle with increasing velocity. Therefore the Q-values become smaller as the temperature increases. The values of cross sections at temperatures T_1 and T_2 are respectively

$$\frac{Q_1}{Q_2} = \frac{n\lambda_2}{n\lambda_1} = \frac{1 + C/T_1}{1 + C/T_2} \qquad \ldots\ldots(33)$$

where C is the Sutherland constant having the values $C_{He} = 80$, $C_{H_2} = 80$, $C_{Ar} = 165$, $C_{air} = 115$. As will be seen from (33) the change in the value of Q

is noticeable only for relatively low temperatures. With increasing temperature, however, a certain portion of the gas is ionized. In this case, several collision cross sections, atoms against atoms, ions, and electrons, come into play. Then equation (31) splits up into several terms

$$\eta = \Sigma \tfrac{1}{3} n_i m_i \bar{c}_i \lambda_i. \qquad \ldots\ldots (31a)$$

The suffix i refers respectively to atoms, ions and electrons. The λ are given by

$$\lambda_a = \frac{1}{n_a Q_{a}^{a} + n_i Q_{a}^{i} + n_e Q_{a}^{e}}.$$

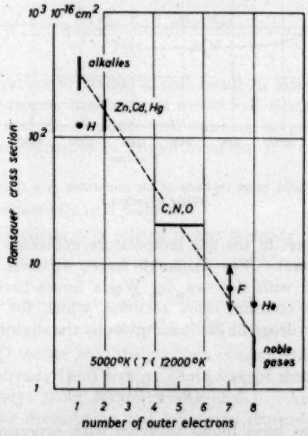


Figure 11. The values of the Ramsauer cross sections for atoms having a different number of outer electrons.

Now the number n_i does not cancel and a pressure-dependent value of η is obtained (the ionization equilibrium depends upon pressure). The Q_{a}^{e} are given by the Ramsauer cross sections. These are approximately $Q_{He} = 6 \times 10^{-16}$ cm^2, $Q_{H_2} = 12 \times 10^{-16}$ cm^2, $Q_{Hr} = 24 \times 10^{-16}$ cm^2 and $Q_{air} = 16 \times 10^{-16}$ cm^2. From the plasma of an electric arc cross sections of atoms for electron collisions have been obtained, as shown in figure 11 (Maecker 1956). It will be shown later that the experimental evaluation of Q makes it convenient to include in Q some small numerical factors of order of magnitude 1. Therefore the results should actually be represented on a slightly different scale. As the accuracy of experimental values is between $\pm 15\%$ and $\pm 25\%$ we may safely neglect this difference here. The values of the Ramsauer cross sections are considerably larger than collision cross sections from kinetic theory.

A detailed study, the principles of which will be reported later (§ 2.2.2), shows that in the case of molecular gases the values of cross sections given are mean values.

The true values increase by a factor of about 2 when the temperature of the gas is just high enough for dissociation of the molecules, the variation being shown in figure 12.

A similar curve is expected for Q values of molecules colliding with ions and likewise for molecular–molecular collisions, though these curves are not yet known in detail.

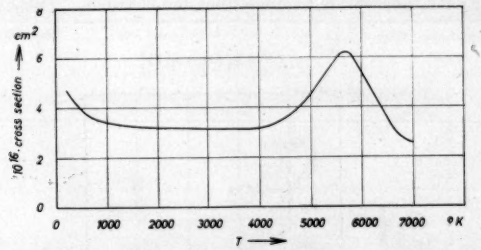


Figure 12. The collision cross sections of air molecules as a function of temperature. (*After* Maecker.)

With further increase in the gas temperature collisions between two charged bodies become important. The Coulomb forces between charges are relatively far-reaching compared with the van der Waals forces between molecules. We therefore obtain large collision cross sections, which for electron–ion collision (ions with charge 1, ' hydrogenic gas ') are given by the slightly modified Gvosdover formula

$$Q_+ = \frac{e^4}{(kT)^2} \ln\left(\frac{kT}{e^2 n_+^{1/3}}\right) \qquad \ldots\ldots(34)$$

which in this form has been found to agree with experiment (Maecker, Peters and Schenk 1955). In this formula e is the electronic charge and n_+ the number density of the ions.

We note that the cross section Q_+ is not dependent upon the nature of the atom. For all singly charged ions at $T = 10\,000°\text{K}$ we find $Q_+ = 2 \times 10^{-14}$ cm².

These collision cross sections being known, it is possible to calculate the dynamical viscosity of the plasma. As an example we try to calculate the viscosity of H_2O vapour at 12 000°K and 1 atm

$$\eta = \frac{n_\text{H} m_\text{H} \bar{c}_\text{H}}{3\left\{n_\text{H} Q_{\text{H}}^{\text{H}\uparrow} + n_\text{O} Q_{\text{H}}^{\text{O}\uparrow} + 2n_\text{i} Q_{\text{H}}^{\text{i}\uparrow}\right\}} + \frac{n_\text{O} m_\text{O} \bar{c}_\text{O}}{3\left\{n_\text{H} Q_{\text{O}}^{\text{H}\uparrow} + n_\text{O} Q_{\text{O}}^{\text{O}\uparrow} + 2n_\text{i} Q_{\text{O}}^{\text{i}\uparrow}\right\}}$$
$$+ \frac{n_\text{i}\left(\frac{8}{3} m_\text{H} \bar{c}_\text{H} + \frac{1}{2} m_\text{O} \bar{c}_\text{O}\right)}{3\left\{2n_\text{i} Q_{\text{i}}^{\text{i}\uparrow} + n_\text{H} Q_{\text{i}}^{\text{H}\uparrow} + n_\text{O} Q_{\text{i}}^{\text{O}\uparrow}\right\}}. \qquad \ldots\ldots(35)$$

The indices H, O, i refer to hydrogen atoms, oxygen atoms and ions respectively, whilst the arrows indicate the collision partners.

Evaluating this equation we put $Q_{\mathrm{H}}^{\mathrm{H}} = 1{\cdot}56 \times 10^{-15}$ cm²; $Q_{\mathrm{H}}^{\mathrm{O}}$ and $Q_{\mathrm{O}}^{\mathrm{O}}$ are not known and might be given approximately the same value:

$$Q_{\mathrm{H}}^{1} \simeq Q_{\mathrm{H}}^{*} = 1{\cdot}3 \times 10^{-14}\ \mathrm{cm}^2,\quad Q_{\mathrm{O}}^{1} \simeq Q_{\mathrm{O}}^{*} = 2 \times 10^{-15}\ \mathrm{cm}^2,\quad Q_{1}^{1} = 2 \times 10^{-14}\ \mathrm{cm}^2.$$

Inserting these values in equation (35) we obtain $\eta = 7{\cdot}0 \times 10^{-4}$ g cm^{-1} sec^{-1}. The viscosity of H_2O vapour at $T = 100°$C is $1{\cdot}3 \times 10^{-4}$ poise. The viscosity therefore, increases by a factor 5 in the temperature interval 400 to 12 000°K. Here, apparently, the long range Coulomb fields play an important part.

When the gas is ionized to about 100%, only singly charged ions being present, we obtain finally

$$\eta = \frac{m_{\mathrm{H}}{}^{+}\bar{c}_{\mathrm{H}}{}^{+}}{3Q_{1}^{1}} + \frac{m_{\mathrm{O}}{}^{+}\bar{c}_{\mathrm{O}}{}^{+}}{3Q_{1}^{1}}.$$

An increase in temperature to 20 000°K will result in almost perfect ionization. Consider a gas like hydrogen, consisting of nuclei and electrons: here no further ionization can occur, and the viscosity will increase because of variation of the Q and $\bar{c}$ values. This variation is proportional to $T^{5/2}$ and is independent of pressure.

1.5.3. *The electrical conductivity of a plasma.*

The electrical conductivity σ is related to the mobility of the electrons

$$\sigma = e n_e b_e; \qquad \ldots\ldots(36)$$

with

$$b_e = e\lambda/mv \qquad \ldots\ldots(37)$$

we obtain

$$\sigma = e^2 n_e \lambda/mv. \qquad \ldots\ldots(38)$$

Putting again $\lambda = 1/nQ$ we see that the problem is connected with the same cross sections which have already been treated above except that the cross sections to be considered are always those for electron collision. As an example we take pure nitrogen gas; the dissociation and ionization depending upon temperature are represented in figure 8. As can be seen from the diagram, at a temperature of 12 000°K the molecules have practically all disappeared whilst the number of free electrons and positive ions is increased to about one seventh of the number of atoms. With further increase of temperature the conductivity of the gas increases considerably. Expressing v in terms of the temperature,

$$v = \sqrt{\frac{8kT}{\pi m}} \qquad \ldots\ldots(39)$$

we obtain

$$\sigma = \frac{e^2 n_e}{\sqrt{(8mkT/\pi)}} \frac{1}{nQ + n_+ Q_+}. \qquad \ldots\ldots(40)$$

n and n_+ depend not only on temperature but also on pressure, as can be seen from Saha's equation. Therefore formula (40) expresses likewise the dependence on pressure.

Inserting the Gvosdover value for Q_+ from formula (34)

$$\sigma = \frac{(kT)^{3/2}}{e^2\sqrt{(8m/\pi)}\ln(kT/e^2 n_+^{1/3})} \frac{1}{1 + (nQ/n_+Q_+)}. \qquad \ldots\ldots(40a)$$

We see from figure 8 that in the temperature range 15 000 to 20 000°K nQ is small compared with $n_+ Q_+$ so that the quotient $nQ/n_+ Q_+$ can be neglected compared with unity. In the limiting case of a *fully ionized hydrogenic plasma* σ is still proportional to the three halves power of the temperature, whilst all the other quantities remain constant. For this case the resistivity is given by the expression

$$\rho = (7{\cdot}6 \times 10^{-2})/T^{3/2}\ \Omega\ \mathrm{cm}^{-1} \qquad \ldots\ldots(41)$$

independent of pressure, where T is in units of 10^4 deg K. Thus even for 5×10^4 deg K, so far the highest sustained temperature, the resistivity of hydrogen gas is about $7 \times 10^{-3}\ \Omega\ \mathrm{cm}^{-1}$, i.e. only about a thousand times the resistivity of Ni. We note further that formula (40) shows no dependence upon gas pressure, as was the case for the viscosity. A conductivity in the gas comparable with metallic conductivity, which is important for phenomena concerning hydromagnetics, is therefore related to much higher hydrogen temperatures still.

Returning to the example given, we see from figure 8 that above 20 000°K double ionization becomes noticeable. For doubly charged ions the Gvosdover cross section Q_+ has to be replaced by $4Q_+$, and in general for z times ionized ions by $z^2 Q_+$. A gas consisting of several components is best described by an effective ionic charge $\bar{z}$ according to the relation

$$\bar{z} = \frac{\Sigma n_z z^2}{n_e} = \frac{\Sigma n_z z^2}{\Sigma n_z z} \qquad \ldots\ldots(42)$$

in which n_z is the number density of ions with charge z. The resulting conductivity of a fully z-fold ionized gas is obtained by the expression

$$\sigma = \frac{(kT)^{3/2}}{e^2 \bar{z} \sqrt{(3m)} \ln\left[(4kT/e^2)\, d\right]} \gamma_E. \qquad \ldots\ldots(43)$$

In this formula γ_E is a factor introduced by the mutual interaction of the charges, with values $\gamma_E = 0{\cdot}5816$ for $z = 1$, $\gamma_E = 0{\cdot}6833$ for $z = 2$, and $\gamma_E = 1$ for $z = \infty$. d is a shielding radius for the activity of the Coulomb forces, dependent on the distance of neighbouring ions $n_+^{-1/3}$, which has been found experimentally to be $d = \frac{1}{4} n_+^{-1/3}$ (Maecker, Peters and Schenk 1955). The electrical conductivity of nitrogen calculated from equation (41) is shown in figure 13.

1.5.4. *The specific heat of a plasma and the thermal conductivity.*

We consider first the enthalpy per gramme of a certain gas;

for atoms we obtain

$$\mathscr{F}_a = \frac{1}{m_a}\left(\tfrac{5}{2}kT + \tfrac{1}{2}D_a + kT^2 \frac{d \ln Z_1}{dT}\right), \qquad \ldots\ldots(44)$$

for ions

$$\mathscr{F}_1 = \frac{1}{m_1}\left(\tfrac{5}{2}kT + \tfrac{1}{2}D_a + E_1 + kT^2 \frac{d \ln Z_1}{dT}\right), \qquad \ldots\ldots(45)$$

for electrons

$$\mathscr{F}_e = \frac{1}{m_e}\left(\tfrac{5}{2}kT\right) \qquad \ldots\ldots(46)$$

where $\mathscr{F}$ is the enthalpy per gramme, D_a the energy of dissociation when there were originally molecules, E_1 the energy of ionization, and $kT^2\, d \ln Z_1/dT$ the

change in the sum of states for the different particles. This term is negligible for low temperatures but is of importance for high temperatures. Taking into account intermediate molecules which may be formed at different temperatures it is possible to show the energy content of a gas as a function of temperature

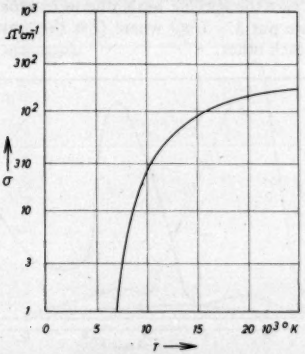


Figure 13. The electrical conductivity of nitrogen as a function of the temperature (atmospheric pressure).

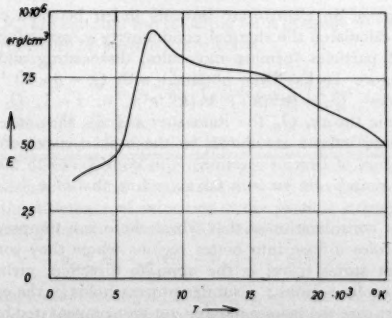


Figure 14. The total internal energy in one cubic centimetre of nitrogen gas as a function of the temperature. (*After* Burhorn.)

(figure 14). The curve obtained is by no means flat which means that the heat of dissociation and ionization is far from being negligible. From the enthalpy we obtain the specific heat C_p of the heated gas at constant pressure p (figure 15).

The thermal conductivity according to kinetic theory is given by the expression

$$\kappa = \frac{n}{3}\bar{v}\lambda C_v \qquad \ldots\ldots(47)$$

which for molecules becomes

$$\kappa = \frac{n}{3}\bar{v}\lambda(\tfrac{5}{2}C_{tr} + C_i). \qquad \ldots\ldots(48)$$

In these formulae κ is the thermal conductivity, C_v the specific heat at constant volume, and C_{tr} and C_i are the specific heats due to external and internal motions of molecules. Again we put $\lambda = 1/nQ$ where Q is the cross section of the corresponding particles for each other.

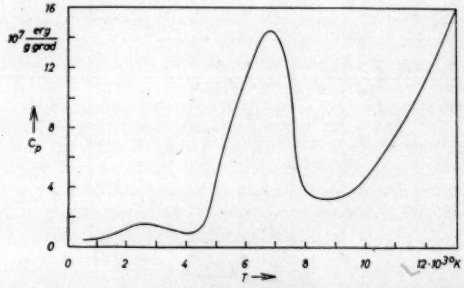


Figure 15. The specific heat of an arc plasma (air and 30% carbon) as a function of the temperature. (*After* Wienecke.)

For the plasma of an electric arc burning in air between carbon electrodes Wienecke (1956) calculated the thermal conductivity κ, assuming the presence of only one kind of particles forming molecules, dissociating and ionizing: The solid curve of figure 16 has been obtained with $Q_c = 8 \times 10^{-16}(1 + 113/T)\,\mathrm{cm}^2$, $Q_R = 20 \times 10^{-16}\ \mathrm{cm}^2$, $Q_+ = \{e^4/(kT)^2\}\ln(kT/e^2 n_+^{1/3})$, $z = 1$, Q_c being the cross section from kinetic theory, Q_R the Ramsauer and Q_+ the ionic cross section.

Very much larger effects are shown by the broken curve of figure 16 which results from diffusion of heat of reaction ; this contributes to the conductivity of heat through a plasma by an amount far exceeding that of κ derived from kinetic theory.

The important consideration is that where there is a temperature gradient in the plasma, molecules diffuse into hotter regions where they consume energy of dissociation, whilst atoms travel in the opposite direction, giving up their heat of dissociation in cooler regions ; a similar process holds in the case of ionization. Now in a stationary case the masses moving out are compensated by masses moving in, the centre of gravity remaining at rest. Each stream of mass-particles $S_m = \rho_m v_m$ carries along a stream of energy

$$W_m = \mathscr{F}_m \rho_m v_m \qquad \ldots\ldots(49)$$

$\mathscr{F}$ being the enthalpy of the mass particle m as before, ρ and v the mass density and velocity of the particle stream. The sum of all particle streams transports the total energy stream

$$W = \sum_m W_m = -\kappa_r \operatorname{grad} T \qquad \ldots\ldots(50)$$

κ_r being the coefficient of heat transfer resulting from reactions. The velocity of the convective mass stream can be derived from the general equation governing motions of a plasma :

$$\rho \frac{dv_m}{dt} = n_m \sum_{m'} n_{m'} \epsilon_{mm'} (v_m - v_{m'}) = F_m \rho_m - \operatorname{grad} p_m - \rho_m \chi_m \frac{\operatorname{grad} T}{T}. \quad \ldots\ldots(51)$$

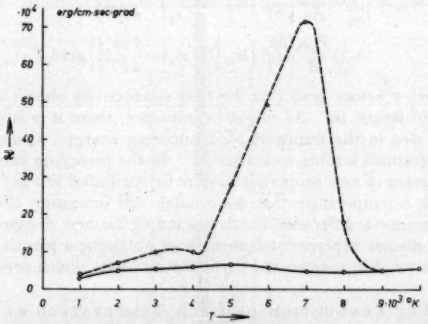


Figure 16. The thermal conductivity of air at different temperatures. Full curve, calculation without, broken curve, calculation with, diffusion of heat of reaction.

In this formula F_m are external forces acting on a gramme of particles of mass m, p_m is the partial pressure of particles of mass m,

$$\epsilon_{mm'} = \epsilon_{m'm} = \frac{8}{3} \sqrt{\left(\frac{2}{\pi} kT \frac{mm'}{m+m'}\right)} Q_{mm'} \quad \ldots\ldots(52)$$

are coefficients of mutual friction.

On the right-hand side of equation (51) we find the forces which drive the stream of mass particles : i.e. external forces, pressure forces and thermal forces. On the left-hand side we find the acceleration of the gas and frictional forces proportional to the difference of velocities of the particle streams.

For the present problem equation (51) reduces to

$$n_m \sum_{m'} n_{m'} \epsilon_{mm'} (v_m - v_{m'}) = - \operatorname{grad} p_m \quad \ldots\ldots(53)$$

with the additional conditions $\Sigma_m S_m = \Sigma_m \rho_m v_m = \sigma$ and $\Sigma_m \operatorname{grad} p_m = \operatorname{grad} p = \sigma$, i.e. that the existing centre of gravity and the existing pressure difference shall not be changed. With three coefficients of mutual friction, ϵ_1 between molecules, ϵ_2 between atoms and molecules and ϵ_3 between atoms and atoms, we obtain the velocities of the respective particle streams :

$$v_m = \frac{1}{A} \left\{ \frac{1}{\rho} \operatorname{grad} p_m \left(m_m - \frac{\epsilon_1}{\epsilon_2} m_a \right) - \frac{\operatorname{grad} p_m}{n_m} \right\} \quad \ldots\ldots(54)$$

$$v_a = \frac{1}{B} \left\{ \frac{1}{\rho} \operatorname{grad} p_a \left(m_m \frac{\epsilon_3}{\epsilon_2} - m_a \right) - \frac{\operatorname{grad} p_i}{n_m} \right\} \quad \ldots\ldots(55)$$

p_m and p_a being the partial pressures of molecules and atoms respectively $p_m + p_a = p \equiv 1{\cdot}016 \times 10^6\,\mathrm{dyn\,cm^{-2}}$; $A = n_m \epsilon_1 + n_a \epsilon_2$; $B = n_m \epsilon_3 + n_a \epsilon_1$. The values of v_m and v_a, evaluated according to formulae (54) and (55), have to be multiplied by their respective ρ and $\mathscr{F}$, the particle density and enthalpy, to obtain the transported energy

$$W = \sum_m W_m = \left\{ \sum_{m_a} \frac{\mathscr{F}_m \rho_m}{A} \left[\frac{\partial p_m}{\partial T} \left(m_m - m_a \frac{\epsilon_1}{\epsilon_2} \right) - \frac{1}{n} \frac{\partial p_a}{\partial T} \right] + \sum_{m_i} \frac{\mathscr{F}_i \rho_i}{B} \left[\frac{\partial p_m}{\partial T} \left(m_m \frac{\epsilon_3}{\epsilon_2} - m_a \right) - \frac{1}{n} \frac{\partial p_a}{\partial T} \right] \right\} \mathrm{grad}\, T. \qquad \ldots\ldots (56)$$

The large bracket before grad T is the heat conductivity shown as a dotted line in the graph of figure 16. As might be expected, there is a large maximum at about 7000°K due to the transport of dissociating energy. Quite similar effects occur at temperatures leading to ionization. In the preceding paragraph we have seen that a number of new properties have to be attributed to a gas which has been heated to such a temperature that dissociation and ionization effects have to be taken into account—a difference which has led to the new designation ' plasma '. In § 2 we will discuss experimental methods of obtaining a plasma and at the end of each section we shall indicate the particular problems which need further study.

§ 2. The Production of High Temperature by Means of the Electric Arc

2.1. *The Arc Technique*

2.1.1. *The free-burning arc.*

The temperature of an arc column is always higher than that of the electrodes, which withdraw heat from the column by thermal conductivity and by radiation. The heat loss is counteracted by surplus energy set free in the region of the cathode and anode fall in front of the corresponding electrode surfaces. The cooling effect due to thermal conductivity is especially noticeable when the electrodes are made of metal. In the column of an arc burning between copper rods in a free atmosphere the temperature is found to be about 4700–5500°K (von Engel 1955); in high pressure mercury vapour lamps the gas temperature is about 5900°K (Elenbaas 1951), whilst the column of an iron arc burning between heat-insulating iron oxide beads reaches 6700°K (Burhorn 1955). This latter value has been obtained from measurements of the Doppler widths of the spectral lines of iron and thus corresponds to the true gas temperature. The free-burning low current arc (about 5 to 10A) between carbon rods reaches a temperature of about 6000 to 7000°K, whilst the electrodes show 3200 to 3600°K at the cathode and 4000°K at the anode. This latter value is related to the temperature of evaporation of pure carbon and has been found to adjust to a constant value of 3995 ± 15°K true temperature, which corresponds to a black-body temperature of 3820°K. This temperature occurs at the maximum value of the arc current just below the hissing point. Several investigations aim at fixing this temperature value for standard use (Euler 1951, 1953 a, b, McPherson 1940). They have shown that the value is quite reproducible for a certain carbon quality, but changes by a small but noticeable amount with slight alterations of the carbon composition.

The presence of any element having low ionization energy in the arc plasma tends to lower the temperature. In an extreme case atoms with high ionization energy are not ionized at all, all the electrons being separated from atoms with lower ionization potential. We would like to draw attention to the fact that in a mixture of gases the resulting ionization is not only dependent upon temperature but also on the concentration of free electrons, as expressed in Saha's equation.

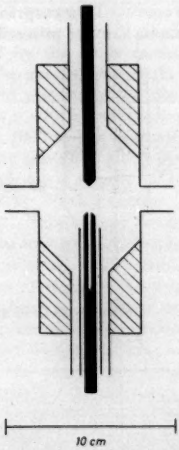


Figure 17. Housing of an electrode stabilized electric arc.

The emission of a free burning arc is subject to sudden violent or small variations, caused by motion of the cathode spot or by gaseous eruptions from the carbon electrodes (Lochte-Holtgreven 1955). These variations are accompanied by variation of the potential drop across the arc. We have tried to steady the arc column by a number of separated water-cooled diaphragms. The diaphragms appeared to reduce the variations both of current and of luminosity to one-fourth of the original value, but the constancy of light emission was nevertheless still unsatisfactory. Good results have been obtained with a housing built according to the design of the Utrecht school (Kruithof 1943 a, b) (see figure 17). Using this device and spectroscopically pure carbon rods, the arc emitted constant intensity to within 5%.

2.1.2. *The low-voltage high-current arc.*

Increase of temperature can be obtained by increase of current. The voltage–current characteristic, which falls for ordinary arcs, rises for high-current arcs. This indicates an increased energy release within the arc. We consider first the low-potential high-current arc burning between a tungsten rod as cathode and

a copper plate as anode, both water cooled (Busz and Finkelnburg 1953, 1954, 1956), see figure 18. This arc, only a few millimetres in length, requires a total potential across the terminals of about 12 v when burning at 500 A. A concentric tube surrounding the cathode is used for gas supply. It has been found that for an arc burning in air or in a helium atmosphere the total potential supplied is less than the ionization potential of the gas. The main portion of the overall voltage is used to sustain the cathode fall of about 30 v cm^{-1}, whilst the field-strength in the column is only 1·5 to 2 v cm^{-1}. This surprising fact demonstrates that in in this arc excitation and ionization are not primarily due to electron impact but

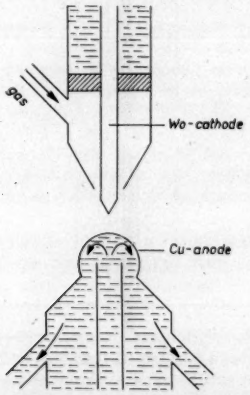


Figure 18. Low-potential high-current arc. *Above*, tungsten cathode surrounded by gas inlet ; *below*, water-cooled Cu anode.

to the high temperature, which measurements have shown to range from 18 000° up to 30 000°K depending upon the composition of the plasma. The total energy fed into the arc is about 6 kw compared with about 600 w dissipated in a low-current arc ; the increased power results in a rise of temperature from 6000 to 30 000°K. It is believed that this low-potential arc is by far the easiest way to obtain high temperatures ; further, the arc can be burnt in any gaseous atmosphere. The pressure of the gas in which the arc is burning might also be changed, which would result in a different energy consumption and consequently in a different temperature.

2.1.3. *The Beck arc.*

H. Beck in 1910–12 showed that a salted anode can stand a considerable increase in current density without hissing of the arc. More specifically, the so-called Beck arc is characterized by an anode containing oxides and fluorides of cerium. In this case the current density at the anode may be several times the maximum normal density. The high current density leads to a very rapid evaporation from

the anode surface and thus an exceedingly large anode drop and a positive slope of the voltage–current characteristic. The maximum temperature measured in this arc is about 9000 to 10 000°K. The possibility of introducing any element into the arc by mixing the substances in question with the filling of the carbon anode seems to be interesting. The rapid evaporation results in a high-temperature vapour jet-stream, which extends primarily in a direction normal to the anode surface, but may be bent into a different direction by the magnetic field of the arc current. The jet-stream has a measured initial velocity of 10^3 to 10^4 cm sec^{-1} and is of particular interest, because there is no further heating by the current, and the luminosity produced by the jet is caused by temperature only. The temperature will drop along the stream, so that the rate of decrease can be calculated from the total length and velocity of the observed jet stream. It is found to correspond to Newton's law ; with a total cooling rate of about 10^{-3} sec to cool the gas from about 10 000 to 3000°K, the decrease of velocity is measured from little irregularities in luminosity travelling along the flame (Finkelnburg 1948a).

The length of the d.c. Beck arc is usually 10 to 30 mm and the total voltage at the terminals about 50 to 60 V. With a current of 200 to 250 A the total energy supplied is thus 12 to 15 kW. This energy is about twice or more than twice the energy required by the low-voltage high-current arc and yet the temperature reached here is only about one third. We conclude therefore that most of the energy is needed to sustain the mechanism of the jet stream with the evaporation of the carbon and to counteract the excellent cooling by external convection. Furthermore the long column creates a considerable magnetic field, the lines of force being directed circularly around the arc column. This field causes any curvature in the direction of the current to increase, thus causing an instability of Beck arcs, which are not straight. When the electrodes of the arc are situated perpendicular to each other the anode flame is bent upwards, partly because of the magnetic field, partly because of particle jets starting from anode and cathode, as will be seen later (§ 5). It is important to realize that this upward bend of the anode flame is not caused by thermal convection, as the arc will show a downward bend when turned upside down. Both from the economic and from the spectroscopic point of view the Beck arc, though interesting, is not really well suited as a high-temperature light source. The latter because each element of volume in the arc changes rather rapidly in chemical composition and temperature. From the two high-current arcs described it is clear that further progress can be expected only by energy concentration into a small volume, and by making conditions inside this volume as stable as possible.

2.1.4. *The Gerdin diaphragm.*

In the years 1922–24 Gerdin (Gerdin and Lotz 1922, 1923, 1924) tried to obtain high-current densities by forcing an arc to burn through a narrow constriction, which was water-cooled to withstand the heat. The temperature reached inside the constriction was higher than expected and reached 20 000 to 30 000°K. This increase in temperature was obtained only in part by increase in current density, the other equally important part being through cooling by the water-cooled diaphragm. This cooling of the outer parts of an arc column withdraws carriers from the conducting plasma. To hold the current at the value it had

without constriction requires a higher voltage, the surplus energy being needed to produce new carriers inside the constriction to maintain the same current value. The carrier production is related to a higher temperature in the centre part of the constricted arc. Therefore the apparent paradox holds that cooling of an arc from without increases the temperature in the central part. The Gerdin diaphragm consisted of heat-resistant earth-like material with a central bore of a few millimetres confining the arc over a length of a few millimetres. Although with the Gerdin diaphragm high current density with an additional heating effect in a stationary state was attained, spectroscopists were not satisfied: most interesting phenomena were expected to occur inside the diaphragm, but observation could only be made obliquely from either side, with a layer of inhomogeneous temperature of equal extent in front of the interesting volume. Therefore the length of the confined portion had to be extended and the inhomogeneous layer had to be made measurable.

2.1.5. *The water-pipe arc.*

In the years 1949–51 Maecker (Maecker 1951, Burhorn and Maecker 1951) tried to confine an arc column by a tube cooled along the inner surface by running water. In this case several effects combine advantageously. A tube cooled by

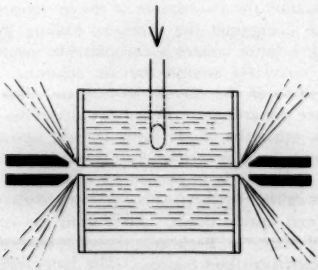


Figure 19. The water-pipe arc. The water inlet is situated tangentially to the arc housing.

running water from inside is not directly exposed to the effect of heat and thus will last a long time; the cooling effect of running water in direct contact with the arc column is of the highest efficiency and, moreover, the tubes can be chosen of small diameter thus confining the arc considerably; the running water can be made to stick to the inner surface of the tube by centrifugal force. Thus a device is obtained as shown in figure 19. The very effective cooling results in a marked increase in voltage along the arc column; potential gradients up to 300 v cm^{-1} have been measured. Again the slope of the voltage–current characteristic is positive, as can be seen from figure 20. With a potential gradient of 240 v cm^{-1} the overall voltage across the arc terminals reaches 500 to 700 v. With a current of 200 A this results in an energy input of 140 kw thereby producing a temperature of 30 000°K. This power input is by no means excessive. By further increase

in current, the temperature can be further raised. The limit to higher temperatures lies in the fact that a further increase in current produces only a relatively slow increase of temperature, as can be seen from figure 21. Quite a number of

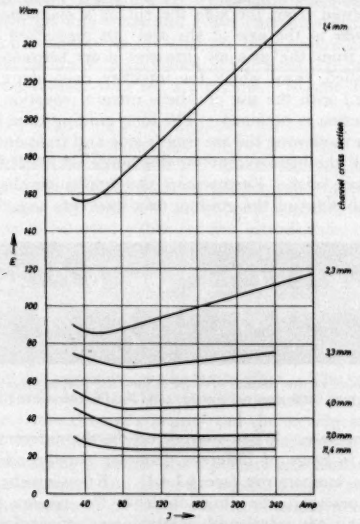


Figure 20. The potential–current characteristic of a water-pipe arc. The parameter is the channel cross section.

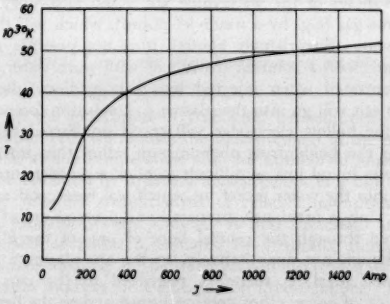


Figure 21. The temperature obtained with a water-pipe arc as a function of current (channel cross section 2·3 mm).

investigations have been made with this arc, the observation of the luminous part being possible either from the side through the surrounding water or else by viewing end-on through a central bore in one of the electrodes. This axial hole in one electrode is also used to guide an extremely narrow carbon rod to ignite the arc. When viewed from the side the slit of a stigmatic spectrograph was situated perpendicular to the axis of the arc. In measuring line intensities for different distances from the axis an intensity drop between different lines of observation is obtained, from which the intensity drop along the radius of the arc can be calculated with the use of Abel's integral equation. Thus the radial temperature distribution is obtained. The long guiding tube allows an easy test for optical thickness in viewing the arc from a side and from one end. As a result it was found that all the lines except for the centre of the H_α line were emitted from an optically thin layer. Furthermore the rapidly developed steam blowing violently out to both sides of the guiding tube prevents any electrode impurities

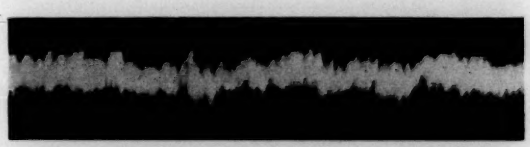


Figure 22. The motion of the column of a water-pipe arc perpendicular to the axis of the pipe (time resolved picture, time from left to right).

from entering the hot zone. When viewed end-on, the different spectral lines will show up strongest in rings of different diameter corresponding exactly to the maximum excitation temperature (see § 1.4.1). Knowing the transition probability and the temperature, the emission from the plasma per centimetre of geometrical depth and per wavelength interval can be calculated, and from this the inhomogeneity in viewing the arc column from the side can be calculated and the measurements can be corrected for a homogeneous layer.

A further advantage of the water-pipe arc is the possibility of adding to the cooling water some gas (e.g. by a water-jet pump), which will then be introduced into the arc plasma. Surprisingly enough it is not possible to introduce salts by running the arc with a solution instead of with pure water. Even when the solution is concentrated when injected into the guiding tube only water will evaporate and no salt will go into the plasma. A solution forced into the arc axis through one of the hollow electrodes will travel unaffected through the hottest part of the arc by the Leidenfrost phenomenon rather than introduce salts to the arc plasma. It was found just as difficult to blow a powder into the arc, because the solid will go into the water jacket by centrifugal force and again keep outside the arc plasma. Only a previously prepared and pressed rod of a solid substance may be introduced through the central bore of one of the electrodes and this will then add, though not quantitatively, to the arc plasma. The light source, although it seemed ideal, is therefore restricted to the use of gases in the vapour of water or of some other cooling liquid and to the limited introduction of solids.

But there is one even greater disadvantage. Investigations of the steadiness of the arc column have revealed that the arc column has by no means a steady position inside the guiding tube, but is blown to and fro perpendicular to the axis by steam development, see figure 22. This latter effect causes serious trouble in the case of narrow guiding tubes and heavy current load, because an element of volume under observation quickly changes temperature and plasma content, allowing only the evaluation of mean values. The trouble is caused by the cooling water being in direct contact with the arc plasma. Recent devices have tried to avoid this.

2.1.6. *Recent development of high temperature arc technique.*

The problem of creating and sustaining temperatures up to 50 000°K is solved by the water-pipe arc. As indicated above, new demands arise as soon as the plasma of the arc is to be used either for the introduction of substances or for precision measurements. Especially for the latter use the water jacket is fatal. Therefore the direct contact of water with the plasma has been avoided and the arc has been burned inside a number of water-cooled copper diaphragms each only a few millimetres in height, and insulated from each other by mica or some other heat-resistant material (Maecker 1956, Lochte-Holtgreven 1956). The central bore of the diaphragms may be made only a few millimetres wide and the current through this assembly may be up to 400 A. The temperature obtained in this way is slightly above 30 000°K. It seems possible to introduce gaseous or powdered substances into the arc's plasma and the steadiness of the arc column has been found to be sufficient for precision measurements. Although the arc technique is still developing it seems that a definite goal has been reached, namely the production, maintenance and use of temperatures up to 35 000°K. Higher temperatures in the region of 50 000°K can be obtained practically, but are for many purposes not yet sufficiently under control.

Before we pass to different methods of obtaining high temperatures we will see what problems have become accessible and have actually been solved with the arc technique described above.

2.2. *The Scientific Use of the High Temperatures obtained by Arcs*

2.2.1. *The evaluation of transition probabilities.*

We turn first to the evaluation of transition probabilities of spectral lines. Starting from equation (10) we require the measurement of the absolute intensity of a line in addition to the temperature and the number of neutral atoms. The temperature can be obtained by intensity measurement of a line of known transition probability or else by measurement of relative intensities of lines whose relative transition probabilities are known. A first successful attack was started by the Utrecht school of research, initiated by Ornstein, Burger, Dorgelo, and others. Here the free-burning arc between carbon electrodes was used, the latter filled with a suitable mixture of elements for investigation. The current was between 5 and 10 A and the whole arc was surrounded by a housing as drawn in figure 17. The temperature obtained ranged between 4000 and 8000°K and

has been evaluated in most cases from the intensity distribution over the rotational structure of CN or C_2 bands. In particular the method of unresolved rotational structure has been worked out in detail and widely used. For temperatures of 5000°K or lower the relative intensities of the copper lines $\lambda = 5153$ and $\lambda = 5700$ Å were used for temperature evaluation, the copper being introduced into the filling of the carbons in such small amounts that self-absorption was avoided. To lower the temperature, elements of low ionization potential, such as the alkaline metals, were added. From intensity measurements some relative transition probabilities were evaluated which were adjusted to the absolute scale by a proportionality factor obtained from one or very few lines, whose absolute transition probabilities were known in other ways. A second attack was made with the help of high current arcs mainly at Kiel. In the water-pipe arc oxygen and hydrogen lines could be measured. As the transition probabilities for the Balmer lines are known, absolute values of oxygen lines could be evaluated. In the newly developed copper-pipe arc air and CO_2 allowed evaluation of N and C lines. Comparison with oxygen lines measured in the same light source gave absolute values. Further increase of temperature has made available lines of He, O^+, N^+ and C^+ whilst He^+ is now being measured as a new 'thermometric line' (see § 1.4.1). The application of arc technique to many more elements depends upon feeding those elements into the plasma at an exact rate. Good progress has been made in this direction, although reliable results have not yet been obtained.

Tables of transition probabilities evaluated experimentally by different methods have been published recently in a number of textbooks and monographs; see, for example, Aller (1953), Biermann (1950), Unsöld (1955). Transition probabilities can be obtained independently by theoretical calculations as shown by Bates and Damgaard (1949) and also by Garstang (1950, 1951). For hydrogen-like terms which can be calculated without interpolation, the calculated values are quite reliable. The same holds true for most terms of more complicated atoms but there are some definite inconsistencies, the cause of which is still unknown.

2.2.2. *The evaluation of collision cross sections.*

The total current is distributed over the cross section of the arc according to the electrical conductivity $\sigma(T)$:

$$I_{\text{total}} = 2\pi E \int_0^R \sigma(T) r\,dr. \qquad \ldots\ldots(57)$$

With

$$\sigma = e n_e b_e;\ b_e = \frac{e\lambda_e}{m\bar{v}};\ \bar{v} = \sqrt{\frac{8kT}{\pi m}};\ \lambda = \frac{1}{nQ + n_+ Q_+} \quad \text{(equations (36) to (40))}$$

we obtain

$$I = \frac{2\pi E e^2}{\sqrt{(8m/\pi)}} \int_0^R \frac{n_e r\,dr}{\sqrt{(kT)}(nQ + n_+ Q_+)}; \qquad \ldots\ldots(58)$$

we insert for Q_+ the Gvosdover cross section, equation (34), and evaluate n_e and T from equations (9*a–d*). The easiest way to find Q is by a graphical solution, inserting some values for Q and drawing an (I, Q) diagram. In this

graph the experimental I value indicates the true Q value. The Ramsauer cross sections obtained in this way are shown in figure 11. We want to make clear that the collision cross sections derived in the way indicated are assumed to be independent of temperature. Therefore the values of figure 11 are mean values for the temperatures present in the arc. Furthermore the collision cross-sections as found above are really ' transportation ' cross sections. We will take the above equations as a definition of Q, and avoid direct comparison of absolute values with Ramsauer's experiments (see § 1.5.2).

For numerical evaluation it should be mentioned that the value of n_e to be inserted in equation (58) has to be fairly accurate and should be measured if possible by several independent methods (see Drawin 1956).

Collision cross sections can be derived independently by theoretical calculations as has been done by Mower (1953), Massey and Moiseiwitsch (1951). In the case of hydrogen, theoretical and experimental values disagree by a factor of about 10, the experimental value being the larger one. Repeated experiments give similar results. The cause of the large deviation is so far not clear.

The value of the Gvosdover cross section has been tested carefully and has been found to agree with experiment but for a small factor (Maecker, Peters and Schenk 1955).

For temperatures between 1000 and 7000°K the dependence of the cross sections upon temperature has been derived in detail for air and is represented in figure 12.

2.2.3. *The evaluation of plasma properties.*

Collision cross sections in a plasma and the correlated phenomena, dynamical viscosity, electrical conductivity, and heat conductivity (§ 1.5) have been evaluated from data obtained with arcs. These plasma properties are really of prime importance for the appearance of the arc phenomenon.

The high current arc is observed to have a strongly accentuated central part. This can be understood from the variation of heat conductivity, which changes markedly with temperature. Heat conductivity transports heat to the outer parts of the arc, where it is removed by convection. As can be seen from figure 16 the heat conductivity for nitrogen has a sharp maximum at 7000°K followed by a sharp minimum at 10 000°K in consequence of dissociation effects. (The evaluation has been described in § 1.5.4.)

A small value of heat conductivity, however, results in a large temperature gradient and consequently in a marked change of radiation. Thus for molecular gases the central part of the high temperature arc contrasts visibly with the rest of the plasma. Similar effects produce the outer corona of an arc : they depend upon the composition of the plasma (Maecker 1953). For atomic gases no such discontinuities are known.

The energy radiated is in many cases so small that it can safely be neglected in the energy balance. However, for the resonance lines and for high pressure discharges, the re-absorption within the plasma becomes noticeable. We refer to the luminous spherical plasma in front of the cathode of a xenon high-pressure arc (Larché 1951, 1952) or to the high-pressure mercury arcs. Here the radiation emitted is a considerable portion of the energy input. In the last case, out of 100 W of energy input 90 W are to be found in radiation, whilst only 10 W are lost

by thermal conductivity. A similar balance holds for the high-pressure water-vapour arc (Peters 1953). Here with a pressure of 1000 atm and a central temperature of 12 000°K the radiation appears to be almost black.

2.3. *The Practical Use of the High Temperatures obtained by Arcs*

2.3.1. *The use of arcs in chemical industry.*

Ordinary low voltage arcs are widely in use for chemical purposes. We remember the technical production of calcium carbide generally performed with three-phase a.c. arcs to heat a mixture of $CaCO_3$ and carbon, or the production of nitric acid by large extended arcs, or the production of artificial sapphires in arcs. However, these technical processes actually require temperatures of only 2000 to 3000°K and the electric arc is used as a means of applying heat in a convenient way to a limited volume.

Temperatures above 4000°K have not yet been effectively used in industrial production. The heat of the arc is sufficient to evaporate any substance. The problem therefore is to obtain a controlled deposit or a synthesis out of the vapour phase. Owing to the great interest in high temperature research, a number of attempts to achieve this have been initiated recently. C_2H_2 is obtained from CH_4 and natural gas by heating the gas with extended arcs followed by 'freezing in' the high temperature chemical equilibrium by means of an effective shower of cooling water. Within the plasma any element is highly active; it therefore generally seems necessary to exclude air and perform the heating and cooling in special protecting gas. Furthermore, once the substance is evaporated there is no further need to keep temperatures high and we shall thus have a high rate of input and output. Hence on these lines it should be possible to make much more universal use of the arc.

As Finkelnburg has mentioned, the high-current Beck arc is most convenient for evaporating any substance introduced into the arc. This can be effected through a central bore in the anode without much consumption of carbon. Although much has to be done to render the preliminary results effective, it appears that this chemistry of the vapour phase opens promising new vistas.

§ 3. High Temperatures obtained with Pulsed Discharges and Sparks

3.1. *Methods for Observation of Pulsed Light Sources*

The high current arcs treated in § 2 had the great advantage of sustaining a high temperature for many seconds, giving ample time for any observations and measurements desired by standard methods. However, we have seen that the energy required to produce temperatures of 50 000°K is already somewhat high and that further progress on these lines can only be expected with much more energy (see figure 21). Further development therefore follows the general trend of impulse techniques in modern physics producing high temperatures only for a short interval of time but with considerably increased power.

The pulsed arc offers a means of increasing the current to several thousand amperes without damaging the fittings of the arc. Pulses may equally well be applied to high or low pressure discharges and permit an increase in the power

input by a considerable factor. As can be seen from § 1.1.2 the time to attain equilibrium decreases with increasing ionization and thus a total pulse duration of milliseconds is ample for quantitative measurements. However, it is necessary to observe only for a fraction of this time in order to be assured of observing a stationary state which does not change during the observation time. The first approach to this has been done by mechanical means using synchronized stroboscopic observation either with rotating mirror or slotted rotating disc. The rotating mirror in combination with lenses can be made to give short duration standing pictures of 10^{-4} to 10^{-6} second quite sufficient for the present purpose (Bartels and Eiselt 1950). A different approach is to use electronic devices. For recorded observation the signal fed to the recording instrument must possess a suitable time constant; this requires the lengthening of the pulse after amplification. The method followed is to observe the pulsed discharge at the right moment by sensitizing the receiving device just during the short interval selected for observation. This can be done by pulsing a photomultiplier, which was closed by a positive potential at one of its amplifying steps. With valves operated by a flip-flop device, the potential can be reversed and the multiplier can be made sensitive for a stipulated period of time. Each pulse will produce a point of the curve to be written by the recording instrument, indicating the energy of the light pulse at that particular wavelength. To obtain the energy profile of a line, one therefore needs quite a number of pulses to scan the wavelength and to obtain the line profile point by point. A prerequisite for a reliable result is thus that successive light pulses be identical. The method is especially appropriate for heavy pulses following each other at long intervals to avoid as far as possible destructive stresses on the materials.

Another method followed by Dieke and his collaborators (Steinhaus, Crosswhite and Dieke 1953) uses a similar device operated at a constant high pulse rate. This method operates at megacycle frequencies and integrates over a large number of pulses. In this way a time-dependent phenomenon repeated at high frequencies can be observed with high resolving power in time for a particular wavelength. The two methods are supplementary with respect to application at low and high frequencies.

3.2. *Temperatures measured in Pulsed Discharges*

Closed discharge tubes may be operated at low power, the current being raised sharply within a short time by a pulse. In this case no difficulties of ignition will arise. The power during the pulse can be taken from a large transformer or from a charged capacitor. A difficulty arises when knowledge of the number of emitting particles is required. The pressure inside the cold discharge tube can easily be measured, but the pressure reached during the pulse cannot be calculated as the temperature distribution inside the discharge tube is generally not known. Pressure readings during the pulse have been made with a Goley cell or a similar pressure-sensitive device of low inertia. However, reliable scaling of the pressure reading is difficult.

There is a further important consideration: the large current during the pulse requires a large electrode surface to avoid troublesome sputtering; but large electrodes require a large volume which is less heated by the discharge.

During the pulse, therefore, most of the gas escapes towards the electrodes and the number density of carriers in the heated part of the discharge tube remains small in spite of a heavy current because of small pressure. Two different approaches are feasible:

(i) the volume at the electrodes is made very large on purpose; then the pressure in the capillary tube remains almost constant during the pulse and the plasma density can be evaluated;

(ii) the electrode space is limited; then a large increase in pressure can be obtained, but the pressure has to be determined by independent measurements. A discharge tube for the latter purpose is drawn in figure 23. We have been able to use a tube as shown in the figure for many thousands of pulses carrying 2000 A and producing a temperature of 30 000°K. Great care should be taken to observe only during constant conditions. As indicated above (§ 1.4.1) some lines show increased intensity with decreasing temperature and thus produce inconsistencies.

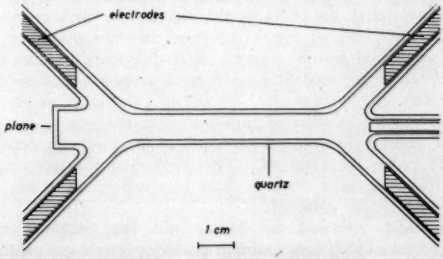


Figure 23. Discharge tube for pulses of 1 to 2×10^3 A. The 'heated' volume is a large fraction of the total volume. Therefore the electron density observed is high although the surface of the electrodes is large, to avoid sputtering. (*After* Wulff 1958.)

3.3. *The Effect of Pressure Waves in Pulsed Discharges*

During observation time a constant temperature is required. It has been found that a number of disturbances in the discharge tube, pressure waves or shock waves, are caused by the sharp rise of current. These disturbances are reflected at the walls of the discharge tube, travel backwards through the discharge and cause additional local heating. Therefore the curve of light emission during the pulse, as obtained by a photomultiplier, shows decrease, increase, or irregular variation of the total intensity emitted, depending upon the shape of the discharge tube and the pressure (Wulff 1958). A typical case is represented in figures 24 (*a*) and (*b*), where the current flow is represented as a function of time. Although it is possible to produce almost square pulses with the aid of valves, the exact square form of the pulse is not necessary and not even desirable from the above considerations. A suitable pulse form can be obtained by connecting in series several capacitors with different time constants of discharge as shown in figure 25. Although the resulting pulse is not very steep at the edges, it is well adapted to spectroscopic investigations and besides it is very simple to obtain. In figure 24 the light intensity emitted is shown as a function of time. As can be seen by comparison, the two

curves are not of similar shape, the effect of pressure waves distorting the curve of light emission. It is quite possible to take advantage of these pressure waves for a further increase in temperature. Necessary conditions are the reproducible production of the pressure waves and geometrically suitable reflecting walls. In a long tube waves will appear in both radial and axial directions. It seems as if the former were more suited for practical use. This leads to a toroidal shaped discharge tube.

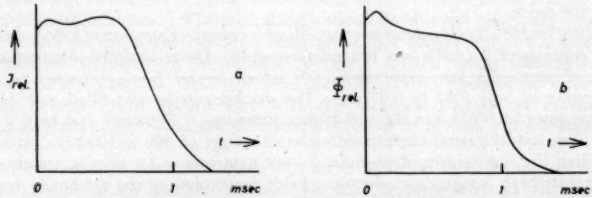


Figure 24. (*a*) The variation of the current, and (*b*) the variation of the integrated light measured during a pulse through the discharge tube (figure 23).

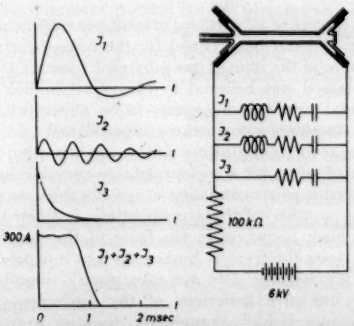


Figure 25. Simple method of obtaining sufficiently ' square ' pulses.

3.4. *The Increase of Temperature by the Pinch Effect*

Besides the temperature increase by pressure waves or more effectively by shock waves there is a further effect resulting in an increase of temperature—the magnetic pinch effect. We need only to remember that parallel currents will attract each other to conclude that heavy discharges are squeezed together by their own magnetic field. This effect is specially interesting, because the contracting discharge is separating itself from the walls of the discharge tube. Therefore the walls are less attacked by the heat and this makes it possible to enclose a very hot plasma in a tube without damaging the latter.

The force exerted by the magnetic field is $\mathbf{j} \times \mathbf{H}$, which is counteracted by the pressure of the confined gas $p = (n_i + n_e)kT$. The theory of the pinch effect has been put forward by Bennett (1934), Tonks (1937) and Tonks and Allis (1939), who have shown that at high currents a conducting column will become closely concentrated at a central axis.

Observation of the pinch effect has been reported by a group of collaborators at Tufts (Fischer 1956 a). Here peak gas discharge currents of 8000 A were used. The temperature was increased by the pinch effect and was found to be $2{\cdot}5 \times 10^5$ deg K.

With 3×10^2 k Joule stored energy a Russian group (Kurtchatov 1956) obtained peak currents of 2×10^6 A in a few microseconds. Quite obviously the capacitors were of unusually low inductance, the whole circuit being reported to have inductance of only 0·02 to 0·03 μH. The discharge tube was filled with H_2, D_2 or rare gases at 0·005 mm Hg and higher pressures. Neutrons and hard x rays were observed, the latter corresponding to an energy of 300 to 400 kev. It seems clear that the mechanism of producing both neutrons and x rays is not thermal. Apparently high longitudinal electrical fields far exceeding the externally applied voltage arise in the discharge tube. The temperature in these discharges has been evaluated from the current following the equation

$$T = I^2/2nk, \qquad \ldots\ldots(59)$$

n being the number density of all particles present per cm^3 including both electrons and ions. Equation (59) can be obtained for the case of thermal equilibrium by expressing the pressure of the ionized gas p by nkT (Spitzer 1956). The value of the temperature obtained was believed to be several million degrees K but the assumption of thermal equilibrium appears to be unjustified. Nevertheless the pinch effect plays a dominant part in these experiments.

Recently a number of investigators followed along the same lines. Using capacitor discharges of some 10^6 A and straight or toroidal discharge tubes, they report having observed a pinch consisting of usually three successive contractions of the plasma. The number of these contractions is limited because the single pinch is unstable. Each curvature in the front of the advancing plasma tends to increase in size; soon the front is broken up into independent sections and a further increase in temperature does not take place. Simultaneously the instabilities perturb the inductive behaviour of the discharge. When the rate of change of the inductance is high enough, the resulting over-voltage can reach a value many times the applied voltage: hence the large electrical fields causing x-ray emission. In this connection it should be mentioned that the magnetic lines of forces stick to the plasma during the pinch, thus increasing the field strength. Compressions of 50 in the first pinch and of 100 in the second pinch with magnetic field strengths increased by the same factor have been observed (Anderson *et al.* 1957).

Theoretical considerations have tried to overcome the pinch instability. It is easy to show that a straight discharge compressed by its own magnetic field leads to an unstable configuration. The instability is of the 'sausage type' ($m = 0$), where the plasma is 'necked' off in one or more places (Taylor 1957). The $m = 1$ mode is a helical or 'wriggle' instability. It follows from the equations

that there are stable configurations. For example, a longitudinal magnetic field having a perpendicular gradient of field strength will lead to a stable configuration for a straight plasma, or in the case of a toroidal-shaped, current-leading plasma an additional external perpendicular magnetic field is required for stability (Biermann and Schlüter 1957).

The temperatures reported (Thonemann *et al.* 1958) are about 1 to 5×10^6 deg K. Numerous investigators report having observed neutrons coming out of deuterium-filled discharge tubes, but in most cases there is some doubt as to the true thermal origin of these neutrons. The critical indications for thermal origin are: (*a*) the number and energy of the observed neutrons should correspond to the numbers and energies calculated for the thermal reaction; (*b*) the energy distribution of the neutrons around the discharge tube should have no preferential direction; (*c*) the neutron yield should vary with the electrical current as calculated; (*d*) the neutron yield should not be strongly affected by a weak external magnetic field.

An interesting dispute has arisen about the influence of impurities of higher atomic weight in the plasma (Allen *et al.* 1957, Anderson *et al.* 1957). The influence is twofold: energy loss by stripping of the electrons, and energy loss by *bremsstrahlung*. The latter seems to be of minor importance, as the time is too short for noticeable energy transfer.

Schlüter (1957) has suggested putting the hot plasma in a periodically changing outer magnetic field. Each cycle of magnetization and demagnetization should throw additional energy into the plasma, thus heating the latter more and more. According to results so far published, these ideas have not yet been tested experimentally.

3.5. *The Temperature Evaluation in Sparks*

As shown in § 1.1.2 a time-dependent equilibrium and thus a temperature can be assumed even for sparks provided the degree of ionization is high enough. An accurate investigation has been made recently by Mandelstam and collaborators (Mandelstam 1956 a, Dolgow and Mandelstam 1953), who measured the expansion and the light emission of the spark channel by a rotating mirror method combined with a Schlieren method. The same spark has been put into an interferometer and from interferometric measurements the gas density and temperature could be derived as a function of time.

By these methods the gas temperature and the excitation temperature could be obtained independently, and it was shown that they agreed within experimental accuracy.

In addition to this a new interpretation of the spark phenomenon was given, in which the spark was considered as a cylindrical shock wave. The front of this shock wave produces the ionization of the gas and the velocity of the front gives the expansion of the spark channel. Knowing the gas pressure from interferometric measurements, the Hugoniot relation could be applied to the cylindrical shock wave, and the values calculated from pure hydrodynamics were found to agree well with the values derived from plasma observations. It seems to the author that this simultaneous approach to the study of the spark from very different directions is one of the necessary prerequisites for further progress.

The temperatures evaluated in these spark experiments ranged from 40 000°K downwards. The time to establish equilibrium between electrons and gas atoms has been shown to be 10^{-7} sec. The main result is that sparks may be readily used as a convenient tool for plasma research. At the same time we see that the problems of sparks are intimately linked with the problems of shock waves, which will be dealt with in §§ 5 and 6.

3.6. *Recent Development of Spark Technique for obtaining High Temperatures*

To obtain very high temperatures in sparks, i.e. high energy dissipation in a short time, two conditions have to be fulfilled: the breakdown time should be as short as possible and the spark resistance during the breakdown should be as large as possible. The first condition requires a circuit of as small inductance as possible. Ideally this inductance can be reduced to that of the capacitor itself. The second condition can be met by the use of high pressure.

Reynolds and Craggs (1942) have tried to obtain a deuterium fusion reaction in sparks. According to Gamow and Critchfield (1949) this reaction requires an ionization of about 100% at a temperature of 10^6 deg K. In the experiment a capacitor bank of 58·5 μF with a stored energy of $1{\cdot}8 \times 10^5$ W was discharged in about 29 μsec with a current of 2×10^5 A. It was hoped that the pinch effect would condense the spark channel, thus further increasing the temperature. A neutron detector was arranged around the spark housing, but no neutrons were detected. It seems desirable to repeat these experiments at various pressures and also to analyse spectroscopically the light emitted.

Further experiments in this direction have been performed by Fischer (1955), using a specially built capacitor of low inductance, which permitted the discharge to occur in a fraction of a microsecond. It was hoped that the energy loss from the spark channel by radiation would remain small during this short interval of time. No special data of the experimental results are known to the author, but apparently a temperature of 10^5 deg K has been reached. Most experiments seem to have been performed in hydrogen or deuterium gas. Although results obtained with these gases are easiest to interpret theoretically, the spectroscopic approach to the highest temperatures might be easier with other gases.

§4. High Temperatures obtained with Exploding Wires

4.1. *The Fundamental Processes related to Wire Explosion*

In § 2 we have seen that an energy dissipation of 6 kW per second and per centimetre arc length results in a temperature of 15 000°K, whilst with 190 kW 45 000°K is obtained (see figure 21). From these values it is concluded that very much more energy must be expended in order to obtain any considerable further increase of temperature. One of the many ways of achieving this was found by Anderson (Anderson 1920, Anderson and Smith 1926), when he discharged a large condenser, charged to 26 kv, through a fine wire approximately 5 cm in length. The energy stored in the condenser was about 30 calories or 125 W sec, which were dissipated in about 10^{-5} second. The energy per second would thus amount to 12 500 kW, about 650 times the energy dissipated in the heaviest arc of figure 21. Thus the method of ' exploding ' wires appears to be one of the most interesting

ways of injecting large amounts of energy into small quantities of matter. And indeed the temperatures obtained would be very high if all the energy were really dissipated in the material of the wire. Much to our disappointment the process of wire explosion, so well known from the burning out of fuses, is more complicated; in the first stage the current warms and melts the thin wire and this is immediately followed by further heating and evaporation. Evaporation of the wire breaks the circuit and it is important to realize that only a certain fraction of the capacitor's energy has been consumed during the first stage of wire explosion. The current interruption by evaporation of the wire is very rapid indeed, and high transient overvoltages can be measured at the terminals of the wire caused by inductance of the circuit or by internal inductance of the capacitor. Only when these inductances are carefully eliminated can the overvoltage be to a certain extent avoided.

A second stage follows the first, i.e. the ignition of an arc which will lead to a complete discharge of the capacitor, usually in the form of an oscillatory discharge. The second stage lasts longer than the first. Both stages are luminous, the first showing a continuous spectrum, whilst the second, after a very bright continuous flash, shows the arc spectrum belonging to the wire material. Only with a circuit of extremely small inductance will emission of some spark lines occur. In the case of very heavy capacitor discharges or very high loading voltages stages one and two of the wire explosion will follow each other in so short a time that the interval of practically no current is no longer observed. Nevertheless, it is believed that the fundamental processes are essentially the same.

The fundamental processes causing this two-stage explosion have been widely discussed (Wrana 1939, Lochte-Holtgreven 1951, Eiselt 1952, Conn 1955). There is no doubt that the sudden evaporation of the material of the wire leads to a very high pressure indeed. After evaporation is completed, the distance between the atoms is still the same as in the solid state. In consequence the material will still be conductive. The following rapid expansion, however, increases the distance between neighbouring atoms. The conductivity drops rapidly. The small mean free paths prevent significant acceleration of the charged particles. Therefore the discharge ceases. It should be understood, however, that the hot plasma always has some small residual conductivity; therefore the current at no time ceases completely.

When the capacitor has been charged to a lower potential and the resulting evaporation is less effective one observes a great number of little sub-arcs all over the length of the wire. Here the quenching of the discharge is obviously due to the fact that the applied potential is not large enough to sustain the many sub-arcs, each requiring electrical potential for anode and cathode drop.

Of especial interest is the re-ignition of the arc. This phenomenon is observed a couple of milliseconds after the circuit break. Doubtless this time interval is related to the expansion of the high-pressure vapour cloud. The re-ignition time depends markedly on the material and also on the diameter of the wire to be evaporated. It was concluded by the author (1951) that the expansion must be sufficiently advanced and the mean free path sufficiently increased to allow the charged particles to take energy from the field. The expansion would lead to a shock wave. Behind the cylindrical shock front there follows a rapid decrease in pressure which favours the re-ignition of the arc. Quite recently this view has

been supported by experimental evidence ; Müller (1957) has obtained excellent short-time photographs of the re-ignition. As can be seen from figure 26, the material of the wire is already expanded, the expansion front showing up as a sharp border. Behind the front, near but not quite at the axis of the expanding cloud, a new discharge develops at a place where the pressure is sufficiently reduced. In a recent report Conn (1955) explains the shock by contraction and subsequent explosion of numerous small metallic droplets, which explode from non-uniform contraction (similar to the ' sputtering ' of droplets from non-uniform

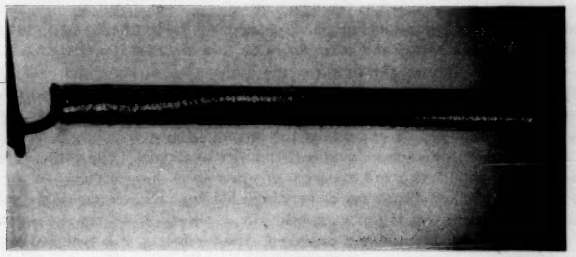


Figure 26.

heat). The shock is assumed to be partially reflected from the vapour cloud surrounding the exploding wire and the concentric cylindrical pressure shock causes the re-ignition of the arc. Although the explanation of the first shock seems to be supported by Schlieren photographs of the shock front (see figure 27), the idea of a partially reflected shock is not correct because Müller's results indicate clearly that the new discharge advances from the electrodes. No trace of a reflected shock could be detected. The spectral lines appearing after the re-ignition are extremely broad and strongly self-reversed by cold vapour surrounding the arc. The self-reversal decreases with time and finally only small emission lines are observed. All this agrees well with the re-ignition inside the expanding cloud.

4.2. *The Evaluation of the Temperatures reached in the Two Stages of the Explosion*

The first stage of the wire explosion leading to continuous radiation is always connected with heated solid material under high pressure. Therefore the optical thickness of the radiating layer is not negligible. On the contrary the radiation may be considered as grey-body radiation. Temperatures evaluated from the energy consumption indicate more than ten thousand degrees.

An attempt to measure the temperature at the beginning of the second stage was made by Anderson himself. He compared the intrinsic intensity of the explosion with the intrinsic brilliancy of the sun. It appeared that about 20 000°K had to be attributed to the flash accompanying the explosion.

Quite recently Bartels applied his method of stigmatic analysis (see § 1.2.1) to the exploding wire. Using quite thin wires he succeeded in producing similar explosions. Arranging the slit of the stigmatic spectrograph at different angles to the axis of the wire he was able to separate the different effects of geometry, optical thickness and absorption. It was shown that the optical thickness in the case under observation was about $\kappa = 1$ and the temperature calculated was $70\,000°\text{K} \pm 10\%$. These measurements were based upon lines showing up above the continuous radiation. As several authors report that the first bright flash of the second stage emits continuous radiation only, the above-mentioned temperature was apparently measured immediately after the maximum intensity of the flash. At higher temperatures the arc lines seemed to be drowned in the continuum, but spark lines appeared instead, which allowed extension of the measurements. So far the limit of applicability for this method of measurement has not yet been reached. It is hoped that this method can be applied even in the temperature range above 100 000°K.

After the bright flash, the temperature drops very rapidly. As practically no spark lines have been observed during the second stage, the temperature in the discharge following the flash cannot be higher than that in a free-burning low-current arc.

4.3. *Further Development of the Exploding Wire Technique for obtaining High Temperatures*

It was pointed out that only a fraction of the energy stored in the condenser is used to heat, melt and evaporate the thin wire. This energy, however, is by no means constant but depends markedly on the rate of energy supply to the wire. Obviously some kind of 'overheating' effect takes place. This rather large effect can be used to obtain higher temperatures. At present, temperatures calculated from the energy consumption in the first stage lie between 10^4 and 10^5 deg K. In spite of this, photographs taken with the rotating mirror technique indicate only an instantaneous, rather weak glow of the wire, followed by a dark phase. Fünfer (1957) has found that the first weak glow lasts for only about a microsecond. He noticed, further, that a wire of solid material heated to 10^5 deg K will cool down through surface radiation and convection according to Newton's law in about a millisecond, i.e. in a time 10^3 times longer than that observed. From this it is concluded that the heat transport in the highly condensed material is so poor that—after an instantaneous radiation—the surface appears dark in spite of the high temperature inside the wire.

This fact is certainly of the highest importance. The rapid energy supply to the thin wire starts some kind of a race between radiation losses and heat conductivity in the material. As the heat conductivity (and also the electrical conductivity) is poor, radiation losses are much smaller than expected. Therefore it would seem that extremely high temperatures can be obtained during the first stage of a wire's explosion without serious radiation losses. This result is the more important, as radiation losses are often prohibitive in high-temperature work. The rate of supply of energy depends on the current rise, which in turn is governed by the total inductance of the discharge circuit. The latter consists of the 'external' inductance of the leads forming the discharge circuit and the 'internal'

inductance of the capacitor. This latter puts a definite limit to the current rise. For this reason it is desirable to develop capacitors having an internal inductance as small as possible. A systematic study of the overheating effect depending

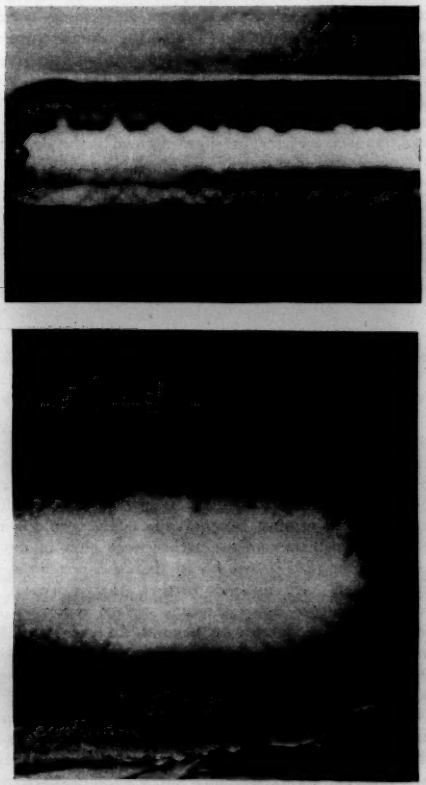


Figure 27. Schlieren photograph of an exploding wire. Two shock waves initiated by stage one and stage two of the explosion can easily be distinguished. (*a*) and (*b*) same explosion at different times.

upon the rate of supply of energy has not been published so far but would doubtless be of great help in estimating the possibilities of further gain in temperature.

Characteristic for the second stage are the shock waves produced. As mentioned in § 4.1, a shock wave is started by the expansion of the evaporated material of the wire. Somewhat later a second, independent shock wave is started with the re-ignition of the discharge. This second shock corresponds to the expanding channel of a spark which causes a shock (see § 3.5) (Mandelstam 1956 a, Dolgow and Mandelstam 1953). These two shock fronts, easily distinguished, can be photographed by a Schlieren method (figure 27).

Anderson has shown already that confinement of the wire in a saw slot made in a wooden block alters the appearance altogether, the continuous light being much enhanced and almost all of the lines appearing in absorption. Confinement of exploding wires leads to the reflection of the shock waves which results in many encounters and superpositions of the shocks. This is the reason why the temperature and hence the light emission of exploding wires increases with confinement. A successful approach for the attainment of higher temperatures, therefore, should begin with a systematic, time- and space-resolved study of the shock waves reflected from various modes of confinement of exploding wires.

§ 5. High Temperatures obtained with Plasma Jets and Shock Waves

5.1. *High Temperatures obtained with Plasma Jets*

5.1.1. *Plasma jets within arcs.*

We have seen that in Beck's procedure the current of an arc between carbon electrodes may be increased to many times the normal maximum without hissing. With this heavy current the anode of the arc evaporates rapidly. Evaporated carbon heated at atmospheric pressure to about 6000°K occupies a volume of about 10^5 times that of solid carbon. About the same holds true for any salts embedded in it. An evaporation of 0·5 mm of anode material per second then corresponds to a vapour jet of 5×10^3 cm sec^{-1} initial velocity (Finkelnburg 1948 b). Actually velocities of about 1·5 to 4×10^3 cm sec^{-1} have been measured from small irregularities in the anode jet. The vapour temperature at the front of the anode is 7000°K; assuming an energy loss from the stream only by radiation and assuming furthermore no change in the diameter of the stream, one finds the temperature of the jet stream at every point to be proportional to the velocity. In a measured example, the temperature changed from 9000°K at the anode front to about 3000°K after a stream length of 9 cm (Finkelnburg 1956).

For the mercury spark similar observations have been made (Finkelnburg 1948 c). With a current density at the anode of 20 000 A cm^{-2} (compared with 300 A cm^{-2} at the anode of the carbon arc) measurements gave a stream velocity of 2×10^5 cm sec^{-1} with an initial temperature of 10 000°K. In the case of mercury there are, in addition, jet streams from the cathode. Again the cause of the jet stream is rapidly evaporated mercury as can be calculated in the same way (Finkelnburg 1948 c). A further mechanism of producing plasma jets in arcs relates to confinement. Whenever an arc is forced to burn through a narrow hole, the increased current density results in an increased magnetic force directed in such a way as to produce further confinement of the arc. As a kind of premature pinch effect for medium currents there is a magnetic pressure directed radially

which results in a pressure release in both axial directions from the confining hole (Maecker 1955). The whole mechanism acts as a kind of pump which sucks gas sideways into a constriction and blows it out in an axial jet after heating the gas and transforming it into a plasma. The energy required for the heating and jet formation is taken from the surplus heat developed in the constriction.

The same observation can be made at the cathode spot of an arc, which acts as a confinement in itself, causing a cathode jet stream, i.e. a plasma particle stream directed perpendicularly away from the cathode surface. The actual pressure as calculated from the Lorentz force

$$\text{grad } p = [\mathbf{jH}] \qquad \ldots\ldots(60)$$

is only a few millimetres of mercury. In a high current arc this cathode jet can readily be observed after short circuiting the arc current for a short time. The initial jet velocity has been measured to be 3×10^4 cm sec^{-1} at the front of the cathode and had dropped to 7×10^3 cm sec^{-1} at a distance of 3 cm from the cathode (Wienecke 1955). Increase of the arc current above 500 A will produce a number of very pronounced plasma jets from the cathode. The velocity of these jets has been found to be approximately 10^6 cm sec^{-1} (Wienecke 1958).

These examples have been given to show that plasma beams may occur inside a plasma. The velocities are of the order of magnitude of several times 10^4 cm sec^{-1} whilst the temperatures are somewhat below the plasma temperature.

5.1.2. *The temperature of plasma jets ejected from arcs.*

The plasma jet shooting out from the Beck arc anode is actually observed outside the arc column. The temperature, however, is decreasing exponentially from 9000 to 3000°K over a few centimetres. The technique of increasing the temperature of an arc by quenching and outer cooling offers a means of obtaining plasma beams of higher temperature outside the arc column. To accomplish this the housing of a high-current water-pipe arc (see figure 19) is closed on one side by a solid homogeneous electrode and on the other by a circular carbon disc containing an axial hole. The diameter of this hole determines the inside diameter of the water jacket and the arc is struck in air between the carbon rod as a cathode and the circular disc. The rapidly developed steam can only escape through the circular hole in the anode ring and shoots out of it forming a hot plasma beam, the temperature of which has been found to be 18 000°K just outside the anode (Weiss 1954). Here the velocity of the beam is some 2×10^4 cm sec^{-1}. The temperature distribution can be seen from figure 28. A careful investigation of this plasma beam, however, shows that it is not entirely free from current. The arc burns from the cathode rod to the anode ring but is prevented from touching that part of the ring which faces the water jacket. The arc therefore sticks to the hole in the carbon ring and is blown outside with the plasma beam. The arc is continually searching for and finding a new and shorter way to the carbon ring only to be blown outside again, thus heating and reheating the plasma beam a short distance outside the arc housing. Experiments to increase both the velocity and the temperature of the plasma jet have shown that there is a limit to both velocity and temperature set by the velocity of sound corresponding to the arc temperature. When this limit is reached by increased current, the arc housing

is broken unless it is constructed of steel as can be seen from the following section. The plasma beam obtained in this way is of fairly high temperature and offers a possibility of investigating a plasma without current.

One of these investigations concerns the appearance of plasma oscillations with or without the additional existence of an outer magnetic field. Experiments have been performed to clear up this point, but no oscillations could be detected. It is assumed that this result is not conclusive as the detector was tuned to a wavelength of 10 cm whilst the oscillations certainly cannot have a wavelength essentially larger than a few millimetres owing to the cross section of the plasma beam.

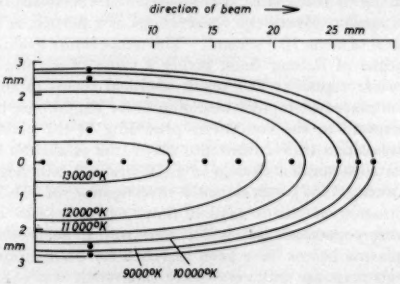


Figure 28. The temperature distribution measured in a plasma jet ejected from a water-pipe arc.

A conclusive experiment would be of extreme value to our understanding of the generation of radio noise in the galaxy. Apart from this a first value for the viscosity of a water vapour plasma at 12 000°K has been derived which, after correction, was $\eta = 9 \times 10^{-4}$ g cm^{-1} sec^{-1} (cf. § 1.5.2). (In the original paper by Weiss (1954) a wrong value for the cross sections was taken.)

5.1.3. *The temperature of plasma jets ejected from arcs with supersonic velocity.*

To obtain plasma beams with supersonic velocity the arc housing has to be constructed of steel. A homogeneous carbon rod at the bottom and a short carbon rod with a central hole at the top serve as cathode and anode respectively. The housing is almost filled with water and the arc struck by a wire inserted in the hole. The heat of the arc evaporates water on a large scale; this emerges from the anode hole as a supersonic plasma beam. The speed of this beam has been measured by Peters (1954 a, b) to correspond to a Mach number 1·5, whilst our own experiments reached a Mach number of 2·6. The temperature measured was about 11 400°K and hence the plasma velocity was of the order of magnitude of 10^6 cm sec^{-1}. The structure of this plasma beam corresponds exactly with the calculations. An interesting experiment was performed by putting an obstacle into the path of the beam. In supersonics the temperature of the reflected beam can be calculated from the temperature of the incident beam by the formula

$$\frac{T'}{T} = \frac{\{\gamma M^2 \sin^2\phi - \frac{1}{2}(\gamma-1)\}\{\frac{1}{2}(\gamma-1)M^2\sin^2\phi + 1\}}{\frac{1}{2}(\gamma+1)M^2\sin^2\phi} \qquad \ldots\ldots(61)$$

where T' is the temperature after reflection and ϕ the angle of reflection, whilst γ is the ratio of the specific heats, which in our case was 1·09. M is the Mach number ; in our case $M = 2{\cdot}6$.

Inserting these values one obtains $T'/T = 1{\cdot}21$ or a temperature rise from 11 400 to 13 800°K. The increase of temperature as determined by experiment was only $T = 1000$°K indicating that a steady shock instead of a reflected shock was observed. This result could be expected, as a reflected shock does not proceed against a supersonic flow. Possibly, however, there is a second reason for the small rise of temperature observed. The beam velocity was 10^6 cm sec^{-1} and the two observations in front and on the surface of the steady shock were made at places 0·06 cm apart. Hence the observations of a particle at the two positions cover a time interval of some 10^{-8} second. The temperature evaluated was obtained from the half-widths of Balmer lines, giving a value of n_e, and from the known composition of water vapour. The result obtained seems to indicate that when an excited atom is placed in rapidly changing outer conditions, the emission does not exactly correspond to the conditions prevailing at the place of observation. It seems that a relaxation time comes into play : the relaxation time required by an atom to adjust to a sudden change in field strength and degree of ionization (temperature) (Döring 1957). Systematic investigation of this effect will allow an exact determination of the relaxation time, which is also of importance in investigations using supersonic beams for aeronautical purposes.

Again these plasma beams have been searched for oscillations. In the 10 cm band there was no response but very likely this result was inconclusive for the reasons already given in § 5.1.2.

5.2. *High Temperatures obtained with Plane Shock Waves*

5.2.1. *Plane shock waves generated by sparks.*

Fowler and his collaborators (1952) used a condensed spark inside a discharge tube to produce shock fronts which could be observed to travel along a straight extension of the discharge tube. The boundary between the strongly luminous expanding gas from the condensed spark and the less luminous gas which it is pushing away is called the contact surface. In front of the contact surface the shock front is advancing. At low pressures the primary shock front is more luminous whilst at higher pressures the contact surface behind the shock is far more luminous.

If the discharge tube is of glass the velocity of the shock can be measured directly by photography with a rotating mirror. In Fowler's experiments velocities of 20 to 40 km sec^{-1} have been obtained which were confirmed by measuring the Doppler effect. From this the Mach number, defined as the ratio of the velocity of the front to the velocity of sound, can be obtained.

Temperatures have been derived from the intensity maximum of lines of Si III, O II, Si II, H, and Si I and were found to vary from 17 500 down to 7000°K along the tube. As the capacitor used for the condensed discharge had 50 μF and a potential of 10 kv, an increase in discharge power and accordingly in temperature is feasible. It is most remarkable that the maximum of electron density as taken from line broadening is not coincident with the maximum of the line of highest ionization potential, Si III, but is somewhat delayed both in space and time. The delay in time is 10^{-6} sec, the delay in space along the shock path about

6 cm, depending upon pressure. The correct meaning of this relaxation time is not clear, as the time required to attain equilibrium should not be as large as this. Perhaps we have here one of the many effects indicating that the line widths do not instantaneously follow a change in external conditions (see also § 5.1.3).

The way to obtain strong shocks and accordingly high temperatures from sparks seems to be of particular interest. Shocks can be created in quick succession without any resetting of diaphragms as is necessary in conventional shock tubes. Sparks can be arranged in a certain configuration and operated in coincidence, thereby producing shock waves of desired configuration. Especially interesting is the magnetic acceleration of a shock already proceeding (Kolb 1956). Probably the highest Mach numbers will be obtained by this procedure.

5.2.2. *Plane shock waves generated in the shock tube.*

The shock tube in its simplest form consists of a cylindrical metal tube containing a high pressure and a low pressure chamber separated from each other by a diaphragm. When the diaphragm is broken a flat shock travels from the high pressure into the low pressure chamber (Turner 1956). The great advantage of this simple device is the easy control of the initial pressures in the two parts of the tube. Using the laws of conservation of mass, momentum and energy the state of the gas behind the shock can be calculated from the known state ahead of the shock and the shock velocity, provided that thermal equilibrium is assumed behind the shock.

The shock velocity can easily be measured either by a number of spark plugs arranged downstream along the shock tube, which will indicate the passing shock front by a change in resistivity over the gap, or else by a number of photomultipliers facing windows in the shock tube wall which will indicate the luminosity of the passing front. Quantitatively the pressure ratio related to the velocity is

$$\frac{p}{p_0} = \frac{(\gamma+1)v_0-(\gamma-1)v}{(\gamma+1)v-(\gamma-1)v_0}, \qquad \ldots\ldots(62)$$

p_0 and p being respectively the pressure before and after the front, v_0 and v the corresponding velocities and $\gamma = c_p/c_v$ the ratio of the specific heats. The temperature on the other hand is obtained from the expression

$$\frac{T}{T_0} = \frac{p}{p_0}\frac{(\gamma-1)p/p_0+(\gamma+1)}{(\gamma+1)p/p_0+(\gamma-1)} = \frac{p}{p_0}\frac{\rho_0}{\rho}. \qquad \ldots\ldots(63)$$

As p/ρ is proportional to the square of the velocity of sound in the gas under consideration a large rise in temperature will occur when the velocity of sound in the moving gas is large and in the stationary gas is small. For this reason hydrogen gas is preferable in the high pressure chamber whilst in the low pressure chamber a heavy gas with a small velocity of sound should be used. In order to avoid dissociation effects which would otherwise consume energy, preferably a heavy rare gas such as argon should be chosen. Formula (63) can safely be used with pressure ratios up to 50, provided there is thermal equilibrium behind the shock. Beyond this value dissociation may occur where molecular gases have been used and accordingly the alteration of the specific heat of the gas has to be considered in formulae (62) and (63).

Experiments using hydrogen and argon gas have been performed by Kantrowitz and his collaborators (Peschek *et al.* 1955, Resler, Lin and Kantrowitz 1952). With Mach numbers ranging from $M = 11{\cdot}2$ to $M = 16{\cdot}8$ temperatures of 10 300 to 13 700°K have been obtained behind the incident shock. Evidently temperatures could be calculated from theory and be observed independently by experiment. These values agreed within the accuracy of the experiment and hence it was concluded that there was temperature equilibrium. As in the experiments referred to in the previous section, the expanding high pressure gas acts like a piston pushing the low pressure gas ahead of it. This creates a shock wave travelling at 4/3 to 9/8 the piston velocity. All the gas between the contact surface and the shock wave has at some time passed through the shock where its enthalpy per gramme was increased to $\mathscr{F} \simeq \frac{1}{2}v^2$, where v is the shock velocity.

In coordinates travelling with the gas behind the incident shock the closed end of the tube may be considered as a moving piston, which creates a reflected shock wave that moves away from the closed end. Across the reflected shock the enthalpy is increased by the kinetic energy of flow of the gas behind the incident shock, i.e. the enthalpy is slightly more than doubled. With Mach numbers ranging from $M = 5$ to $M = 12{\cdot}5$ in the reflected shock, temperatures of $T = 5700$ to 17 000°K have been measured.

Especially interesting are the time resolved spectra taken with a resolving power of 1 μsec. It may be noticed that, due to the gas velocity behind the shock, particles that are seen at times dt apart in laboratory coordinates have passed through the shock wave at times dt' apart, where $dt' = (\rho/\rho')\,dt$. From this it was concluded that with a resolving power of 1 μsec only times of 4 to 8 μsec in particle history could be resolved. The line broadening and shift quickly reach a maximum behind the onset of luminosity. At the lowest Mach number, however, a suggestion of a noticeable increase in ionization with time could be seen on the film, which again indicates a relaxation time of adjustment to the new surroundings. The line broadening and shift was compared with Holtsmark's theory without consideration of the electrons and the agreement was poor. On the other hand agreement with Unsöld's theory of continuous radiation could be obtained after a correction had been applied for the temperature due to cooling during the observation time and after putting the effective charge for argon $z_{\text{eff}} = 1$.

Another interesting point is the heat loss by radiation ; assuming the shock as ' a steady one-dimensional constant area channel flow ' one obtains the relation

$$\frac{1}{I_\nu}\frac{dI_\nu}{dt} = \frac{2}{n_e}\frac{dn_e}{dt}. \qquad \ldots\ldots(64)$$

The quantity on the left-hand side was measured from spectrophotometer oscillograms and that on the right could be determined from line shift measurements. The results agreed in order of magnitude after a 1500°K correction for cooling had been applied.

Measurements of the electrical conductivity (Lin, Resler and Kantrowitz 1955) indicated that for temperatures greater than 11 000°K ionization equilibrium in argon is reached in less than 12 μsec from the time it becomes luminous.

From the above it is seen that shock tube measurements are an unusually powerful tool for elucidating a great number of fundamental processes and relaxation

times. It is confidently hoped that a better understanding may be obtained in the near future by many more new experiments. The number of shock tubes in action is rapidly increasing and many reports in this field give up-to-date surveys of recent experiments.

It should be mentioned, however, that much higher temperatures than those referred to above have actually been obtained in specially designed shock tubes. These include pre-heated shock tubes, shock tubes run with explosions and shock tubes with changing diameters. In the latter case high Mach numbers have been obtained with the use of a large diameter high pressure chamber and a small diameter low pressure region, employing either one or several diaphragms.

5.3. *Further Development to obtain Higher Temperatures with Shock Waves*

5.3.1. *Shock waves of different geometrical form* (*see also* §§ 6.1.5, 6.2.1, 6.2.2).

The shock waves referred to so far have been waves travelling in a tube. It has been shown by Palmer that in this case shocks consist of a thin annulus of cool gas (Chabai and Emrich 1955) which surrounds the main body of the gas in the shock interior the properties of which are quite uniform (Palmer 1956). Again, it was confirmed that these properties agree well with those calculated, using the shock-front velocity as the measured parameter.

Many subsequent researches have dealt with the subject of curved shock fronts. There may be point sources for shocks, produced by sparks, or shocks produced by spherical or other pistons. Specially interesting is the idea of Schmidt (1955) who proposes a repeated spark placed at one focus inside an ellipsoidal body. It is expected that the shock front reflected from the walls will focus at the second focal point and then return to the first focus, at which instant a second spark is initiated. Since the spark frequency is in resonance with the return time of the shock front an accumulation of shock energy is expected. This experiment, although not yet performed to the author's knowledge, is certainly most interesting. However, the desired temperature of 10^8 deg K probably cannot be reached in this way for the following reason: the density at 10^6 deg K will be low and the mean free path correspondingly large. In order to have a 'volume' heated to 10^6 deg K this volume should have an extension of several mean free path lengths. This makes the 'focus' rather large. If the reflecting walls are of solid material they can stand only perhaps 10^3 deg K. A shock travelling from the surface of a sphere with radius r towards the centre causes a temperature rise proportional to $1/r^{0.8}$ (Guderley 1942). A temperature rise of 10^3 necessitates a compression ratio $r_a:r_i$ of 10^3. This means the containing vessel should be very large indeed, the wave expanding from an extended focus cooling down to 10^3 deg K at the walls and being reheated by compression. It looks as if dimensions will not permit us to obtain the high temperatures desired for nuclear reaction in this way. However, in any case the performance of the experiment should be interesting.

Cylindrical shock waves created by the pinch effect and also by the sudden heating are responsible for the high temperatures apparently obtained by Kurchatov (lecture at A.E.R.E., Harwell) whose experiments have been discussed in § 3.4. Although it is not certain that temperatures of more than 10^6 deg K have actually been obtained the influence of colliding shock waves is of prime importance. Systematic experiments in this direction are highly desirable.

ADDENDUM (§ 6)

HIGH TEMPERATURES OBTAINED WITH CHEMICAL REACTIONS

6.1 BY R. SCHALL†

6.2 BY F. WECKEN†

† Attaché to the Services Techniques de l'Armée Française.

§6. HIGH TEMPERATURES OBTAINED WITH CHEMICAL REACTIONS

6.1. *High Temperatures obtained with Detonations*

As energy release by chemical reactions is based on the formation of molecules (like CO_2, H_2O, etc.) which are not stable at temperatures above 6000°K, product gases cannot reach higher average temperatures. Nevertheless, especially with detonating high explosives, higher temperatures can be obtained either by local energy concentration or by energy transfer from a large amount of product gas to particles of smaller heat capacity. The full understanding of these processes implies a detailed knowledge of detonation phenomena which will be treated in §§ 6.1.1 to 6.1.5 (Taylor 1952, Eyring *et al.* 1949, Finkelstein and Gamow 1947, Lebeau and Trombe 1950).

6.1.1. *The three mechanisms of reaction: combustion, explosion, detonation.*

The energy content of explosives is not extremely high. On the contrary, the oldest way used by mankind to produce heat, i.e. the burning of wood in open air, furnishes more energy (4000 cal g^{-1}) than modern military high explosives (1400 cal g^{-1}). The technical importance of the latter is due to the fast energy release. For example, a stick of PETN of 1 cm^2 section, detonating at a loading density $\rho_0 = 1{\cdot}7$ g cm^{-3} with a detonation velocity $D = 8000$ m sec^{-1}, releases $7{\cdot}5 \times 10^8$ m kg sec^{-1} or about 10^7 h.p.

The temperatures which product gases reach after reaction can easily be calculated under adiabatic conditions, provided the reaction heat Q, the specific heat c_v (as a function of pressure and temperature) and the final state of the gas are known. Three cases are usually considered: burning at constant pressure, explosion at constant volume and detonation.

The resulting pressure increases in this sequence and accordingly temperatures rise. Homogeneous reaction temperatures at constant volume are given by the equation $T = Q/\bar{c}_v$ (sometimes called the 'adiabatic explosion temperature'). For typical high energy reactions in gases they are of the order of 3000°K.

In practice the decomposition of an explosive in a closed vessel is never homogeneous. The reaction may begin at any point and the gas expanding from this point compresses and heats the unreacted explosive. By this the reaction is accelerated, and 'explosion' occurs during which energy from reacted particles is transmitted to unreacted ones. This energy flow increases with reaction speed, but is limited by the energy transfer rate of sound waves.

Highest reaction speed is reached when the pressure waves are steepened to a shock igniting the reaction.[2] This mechanism is called 'detonation'. The highest shock velocity sustained by the reaction equals the energy transfer speed, being the sum of sound and material flow velocity. The identity of this sum and the detonation velocity is the content of the Chapman–Jouguet (C–J) rule.

6.1.2. *The hydrodynamic theory of the plane detonation wave.*

The state of the detonating gases after reaction depends on the wave geometry. The simplest case is the one-dimensional wave.

It is convenient to consider a stationary plane wave in a system moving with the detonation velocity D. The continuity of the flows through two reference planes of unit area, one before (index 0), the other behind the shock (without index), leads to the following equations:

$$\rho_0 D = \rho(D-W) \qquad \text{(mass)} \qquad \ldots\ldots(65a)$$

$$p_0+\rho_0 D^2 = p+\rho(D-W)^2 \qquad \text{(momentum)} \qquad \ldots\ldots(65b)$$

$$\Delta Q+\tfrac{1}{2}D^2+p_0 v_0 = \Delta E+\tfrac{1}{2}(D-W)^2+pv \quad \text{(energy)} \qquad \ldots\ldots(65c)$$

where $\rho \equiv 1/v$ = density, W = material velocity, p = pressure, Q = reaction heat per unit mass, and E = internal energy per unit mass.

After the reaction, the state of the gas may be calculated by adding the C–J rule:

$$D = a+W \quad (a = \text{sound velocity in the reacted gas}) \qquad \ldots\ldots(65d)$$

and

$$p = p(\rho, a), \quad \text{the equation of state of the product gas.} \qquad \ldots\ldots(65e)$$

The solution of (65*a–e*) can easily be found for perfect gases and Abel gases with constant covolume. Solutions and numerical values for the quantities in the plane of completed reaction (C–J plane) are compiled in table 2, (*a*) for a typical gaseous explosive and (*b*) for a representative condensed high explosive.

Table 2. State of the Gas behind a Plane Detonation Wave

(*a*) Gaseous explosive $p/\rho = RT$		(*b*) Condensed explosive $p(1-\alpha\rho_0)/\rho = RT$	
$\rho_0 = 10^{-3}$ g cm^{-3}		$\rho_0 = 1{\cdot}7$ g cm^{-3}	
$Q = 1400$ cal g^{-1}		$Q = 1400$ cal g^{-1}	
$\gamma = 1{\cdot}4$		$\gamma = 1{\cdot}25$	
$\bar{c}_v = 0{\cdot}33$ cal g^{-1} deg^{-1}		$\bar{c}_v = 0{\cdot}33$ cal g^{-1} deg^{-1}	
		$\alpha = 0{\cdot}4$ cm^3 g^{-1}	
Equation	Numerical value	Equation	Numerical value
$\rho/\rho_0 = \dfrac{\gamma+1}{\gamma}$	1·7	$\rho/\rho_0 = \dfrac{\gamma+1}{\gamma}$	1·17
$p = 2\rho_0 Q(\gamma-1)$	50 atm	$p = 2\rho_0 Q\dfrac{\gamma-1}{1-\alpha\rho_0}$	170 000 atm
$T = \dfrac{Q}{\bar{c}_v}\dfrac{2\gamma}{\gamma+1}$	5200°K	$T = \dfrac{Q}{\bar{c}_v}\dfrac{2\gamma}{\gamma+1}$	5000°K
$D = [2Q(\gamma^2-1)]^{1/2}$	3400 m sec^{-1}	$D = \dfrac{1}{1-\alpha\rho_0}[2Q(\gamma^2-1)]^{1/2}$	8000 m sec^{-1}
$W = D\dfrac{1}{\gamma+1}$	1400 m sec^{-1}	$W = D\dfrac{1-\alpha\rho_0}{\gamma+1}$	1200 m sec^{-1}

From this table it can be seen that for gaseous and condensed explosives, detonation temperatures are increased by a factor $2\gamma/(\gamma+1)$, with respect to the adiabatic explosion temperatures. This gives a gain of 10–20%.

In the case of condensed explosives detonation pressures of 100 000 atm or more are produced. For more precise calculations in this pressure range the

variation of the co-volume α must be taken into account. Many attempts have been made to find appropriate equations of state consistent with high pressure physics which describe the measured increase of detonation velocity D with loading density ρ_0. The $D(\rho_0)$ curve gives an expression for the behaviour of the real gas, whilst for perfect gases D is independent of ρ_0.

Of the different forms of equations of state used for precise calculations, two limiting cases will be discussed: the variation of co-volume as a function of volume only, $\alpha = \alpha(v)$, and as a function of pressure, $\alpha = \alpha(p)$.

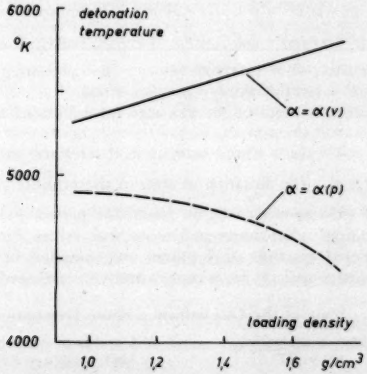


Figure 29. The theoretical detonation temperatures for PETN derived for different co-volume models.

The first assumption implies that the internal energy depends only on temperature, while for $\alpha = \alpha(p)$ pressure contributes to internal energy in the form of intermolecular potential energy. The difference in the distribution of internal energy in the two models has a characteristic consequence: in the first model the detonation temperature increases monotonically with increasing loading density, while in the second it decreases. Figure 29 shows the different slope of the temperature–loading density curve obtained with $\alpha = \alpha(v)$ (Cook 1947) and with $\alpha = \alpha(p)$ (Murgai 1953) respectively. In spite of the difference in temperature, the equations result in the same values of detonation pressures and gas densities, i.e. in the same total internal energy.

6.1.3. *Spectroscopic measurements of detonation temperatures.*

The uncertainty of the theoretical results makes it necessary to determine detonation temperatures by experimental methods. As temperatures are high, only spectroscopic technique can be applied. Fox (1945) has studied spectroscopically the light emitted from the outer surface of the explosive detonating in a non-luminous medium. The spectrum was projected through a grating on to a film in a rotating-drum camera.

All typical high explosives show continuous spectra. Except for dyed explosives their intensity distributions correspond to that of a black body. Colour temperatures measured by Fox are in close agreement with calculated adiabatic explosion temperatures. However, dyeing the explosive had a pronounced effect on the colour temperature.

This result indicates that the photographed light is emitted from expanding gas particles and filtered by unreacted explosive particles. Recently (Gibson *et al.* 1957) a technique has been developed which makes it possible to study the light emitted from the core of the exploding stick. To do this, a rod of transparent material (methylmethacrylate) is embedded in the stick and through this the light is observed. Colour temperatures determined by this technique are as follows :

	$\rho_0 = 1{\cdot}0$ g cm^{-3}	$\rho_0 = 1{\cdot}6$ g cm^{-3}
PETN	5250°K	5750°K
TETRYL	4750	5400
RDX	5800	5600
EDNA	5500	5250

It is most interesting that for some explosives the temperatures increase with increasing loading density ρ_0 while for others they decrease, just as would be expected from the two theories mentioned above.

6.1.4. *Temperatures of the detonation front.*

The thermodynamic quantities given in table 2 are valid for the state of the gas after the completed reaction. Now we will consider the reaction zone confined between the shock front and the C–J plane.

The equations (65*a–e*) can be represented graphically. From (65*a* and *b*) which are valid for any plane behind the shock, one obtains the equation of state

$$p - p_0 = -(\rho_0 D)^2 v + \rho_0 D^2. \qquad \ldots\ldots(66)$$

This means that in a (p, v) diagram equation (66) is represented by a straight line ($P_0 P_s$, figure 30). The slope of this line is determined by the C–J condition (65*d*) which in this diagram makes the state line (65*e*) a tangent to the Hugoniot curve of the product gas.† The tangent point P_{CJ} characterizes consequently the behaviour of the reacted gas. The conditions at the shock front are represented by the intersection point P_s of the state line with the Hugoniot curve of the intact explosive. It can be seen from the figure that the front pressure p_s is higher than the detonation pressure p_{CJ}. Theory gives a ratio $p_s/p_{CJ} = 2$ for both perfect (Döring 1943) and Abel (Eyring *et al.* 1949) gases. This brings the shock pressure of high explosives up to 400 000 atm, a value which has been confirmed experimentally (Duff and Houston 1955). The pressure within the reacting zone decreases with the distance from the shock.

We are now interested in the temperatures reached in the detonation front. The change of internal energy can easily be calculated from (65*a–c*). For $p = p_s$, $Q = 0$ there results

$$\Delta E = \tfrac{1}{2}(p_s + p_0)(v_0 - v_s). \qquad \ldots\ldots(67)$$

† The Hugoniot curve represents all (p,v) states resulting from plane shocks.

As the specific heat $\bar{c}_v$ is defined by

$$\Delta E = \bar{c}_v \Delta T \qquad \ldots\ldots(68)$$

it is necessary to have high shock pressure and great compressibility in order to obtain high temperatures. For homogeneous explosives these two conditions cannot be satisfied simultaneously. For gaseous explosives the shock pressure is low, for condensed explosives the compressibility is poor. Front temperatures of only about 1000°K can be expected for both.

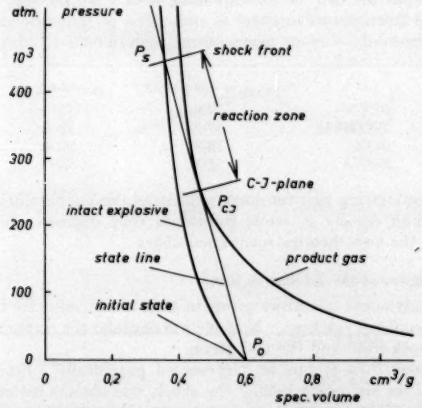


Figure 30. (p, v) diagram of a plane detonation in a condensed homogeneous explosive.

Consider now a condensed explosive containing small quantities of included gas forming pockets of dimensions which are small compared with the thickness of the reaction zone.† In practice this case of high detonation pressure and high compressibility is realized by a liquid explosive containing gas bubbles or by a granular explosive with air interstices. Assuming uniform pressure for the gas and the condensed phase Paterson (1955) calculated shock temperatures of the order of 10^6 deg K for air pockets in PETN at atmospheric pressure. The actual pressure distribution, however, will not be uniform.

The compression of the gas is connected with a rarefaction in the ambient explosive. Pressure equilibrium will be established only after a certain time by pressure waves from more distant regions. This justifies Bowden's treatment of the phenomenon (Bowden and Yoffé 1952) as an adiabatic compression from which pocket temperatures of the order of 100 000°K result. These temperatures are so high that radiation losses will be important in the energy balance. Though there is actually no theory which takes into account the radiation loss, which

† For military explosives the thickness of the reaction zone has been found to be of the order of 1 mm.

lowers the temperature obtained, two facts may be pointed out: (*a*) because of the compression a local energy concentration in the gas pockets takes place, and (*b*) in the gas pockets peak temperatures far above reaction equilibrium temperatures are obtained. These will increase with reduced initial pressure and heat capacity of the enclosed gas.

There is practical evidence that air pocket effects are important for initiation, wave propagation, and also for detonation luminosity (Paterson 1955).

Apart from temperature equilibrium radiation, a detonating high explosive emits electromagnetic waves of nearly all wavelengths. Emission of short radio waves (Campbell *et al.* 1955) and the reflection of these at the detonation front (Koch 1953) has been used as an argument for extremely high detonation temperatures (Sänger 1954). However, it can be proved that this effect is rather due to accelerations of ions, especially in the turbulent flow. This view has been confirmed in the case of aluminized explosives where secondary reaction of the expanding gas with the ambient air gives a second emission peak.

In a short note Fonberg (1951) reports having detected x-rays emitted from a detonation of explosives with high hydrogen content. This exciting statement has neither been confirmed nor rejected. It is hard to understand the occurrence of particle velocities of some 100 v which would seem to be the minimum to yield a detectable radiation.

6.1.5. *Converging detonations.*

We have seen that in the reaction zone an energy flow from the reacted gas to the shock front takes place. In converging detonation fronts, where the shock surface is smaller than the C–J surface, higher energy densities can be expected than for plane detonation waves. Front curvature r and thickness of the reacting zone l are the controlling factors. Theory indicates (Penney 1950) that a cylindrical ($n = 1$) or spherical ($n = 2$) converging detonation in a condensed explosive behaves approximately like a plane detonation wave, the loading density of which is increased by a factor $1 + nWl/Dr$.

Therefore by means of converging detonations states can be realized which cannot be obtained by plane waves.

Though in principle it is not difficult to produce spherical converging waves using explosive lenses as first proposed by Demougin (1930), there are serious obstacles to the optical observation of the centre. Further efforts in this field have not been reported so far. The production and observation of cylindrical shock waves is far easier. Typical effects can be studied in this two-dimensional case although the energy concentration is lower.

Some uncertainty exists about the stability of converging waves. From the theoretical point of view (Butler 1955) converging shocks are unstable. Small perturbations in the front grow during the contraction in such a manner that the ratio of perturbation curvature to mean curvature tends to infinity when the wave reaches the centre. In practice there is no evidence of such instabilities either by flash radiographs of cylindrically converging shocks (Schall and Thomer 1951) or by photographs of colliding detonation waves (Fünfer and Müller 1954). It seems that in these experiments shock strength was not sufficiently high to let instabilities become effective.

Figure 31 shows an experimental arrangement used to photograph spectra near the centre of convergence. Such a spectrum taken without time resolution shows an increase of total intensity near the centre, but the intensity distribution—i.e. the colour temperature—is nearly the same in the centre as in the outer regions or in a single plane wave.

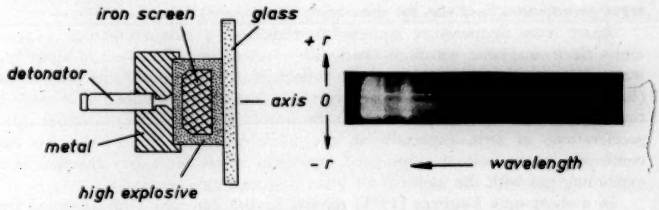


Figure 31. Experimental arrangement and spectrum of a colliding detonation.

This result is not surprising as we have seen that converging waves behave like plane waves of higher loading density and the colour temperature varies little with the loading density. Nevertheless the mechanical effect produced on an iron plate indicates a strong energy concentration in the centre. In the case of condensed explosives, however, this energy is mostly taken from interatomic potential energy and not from thermal energy.

6.2. *High Temperatures obtained with Shock Waves from Detonations*

6.2.1. *Shock-wave generation by detonation.*

In the following sections the transfer of explosive energy from the product gases to an inert gas surrounding the high explosive will be considered. Let a plane detonation wave in a solid explosive progressing in positive x direction reach the boundary $x = 0$ at the time $t = 0$; suppose the half-space $x > 0$ be occupied by an inert gas. Neglecting the extension of the reaction zone we can say that at the time $t = 0$ the gas (with $p = p_0$, $T = T_0$, $\rho = \rho_0$, $u = 0$) borders on the product gas in the C–J state ($p = p_1 = p_{CJ}$, $u = u_1 = W$). A shock wave starts in the positive x direction, a rarefaction wave in the opposite direction. The properties of these two waves are determined (Döring and Burkhardt 1944) in a similar way as in the shock tube. Consider the (u, p) diagram (figure 32). A descending curve A is correlated to the isentropic rarefaction wave in the product gas and an ascending curve B is characteristic for a shock wave in the surrounding gas. The point of intersection of these two characteristic curves gives the values $p = p_2$ and $u = u_2$ for the region in between the two separating waves. From this the 'shock strength' p_2/p_0 can be evaluated. If for example an Abelian equation of state is supposed to hold for the product gas and if the inert gas is assumed to be perfect, the equations of the two curves are

$$\text{Curve A} \qquad u = u_1(\gamma_1 - 1)^{-1}\,[3\gamma_1 - 1 - 2\gamma_1(p/p_1)^{(\gamma_1-1)/2\gamma_1}] \qquad \ldots\ldots(69a)$$

$$\text{Curve B} \qquad u = (2/\rho_0)^{1/2}\,(p - p_0)\,[(\gamma_0 + 1)p + (\gamma_0 - 1)p_0]^{-1/2}. \qquad \ldots\ldots(69b)$$

With p_2 the values ρ_2 and T_2 behind the shock as well as the shock wave velocity U are fixed, and we have

$$T_2 = T_0 \frac{p_2(\gamma_0-1)p_2+(\gamma_0+1)p_0}{p_0(\gamma_0+1)p_2+(\gamma_0-1)p_0}. \qquad \ldots\ldots(70)$$

Generally the adiabatic curve of the product gas and consequently the curve A is not known. In table 3 some values of p_2, p_0, U, u_2 and T_2 taken from the literature are listed. These values have been calculated using the experimental data indicated

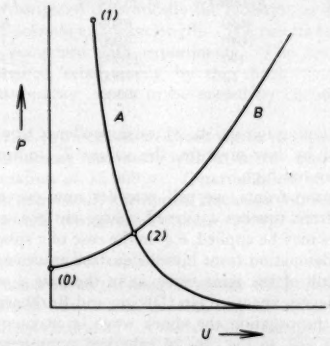


Figure 32. Determination of the shock strength.

in the last column by means of a theoretical equation of state of air (Döring and Burkhardt 1944). This equation is not quite correct because the dissociation energy of nitrogen used was too low (Christian *et al.* 1955).

Table 3. Initial Values of Shock Wave

Explosive	Gas	p_0 (atm)	p_2 (atm)	u_2 (m sec^{-1})	U (m sec^{-1})	T_2 (°K)	Measured	
RDX	air	1	1100	9100	10000	15000	U	(1)
PETN	air	1	1100	9100	10000	15000	U	(1)
Lead azide	air	1	510	6100	6700	8300	U	(1)
TNT	air	1	430	5600	6200	7800	U	(2)
TNT	air	4	800	3700	4200	5500	U	(2)
Tetranitro-methane + toluene	air	1	760	7600	8300	10500	T_2	(3)
	argon	1				27000	T_2	(3)
PETN + TNT	air	1	1040	8600	9500	13100	U	(4)
RDX + TNT	air	1	740	7500	8200	10000	U_2	(5)

References: (1) Savitt and Stresau (1954); (2) Weibull (1947); (3) Stettbacher (1948); (4) Johansson and Ljungberg (1953); (5) Schall (private communication).

For a given explosive the shock-wave temperature T_2 depends on p_0, T_0, M (molecular weight) and C_v (molar heat capacity) as follows: T_2 increases with T_0, with M, with p_0^{-1}, and with C_v^{-1}. Highest temperatures are therefore obtained with heavy rare gases of diminished pressure. The intense radiation emitted

thereby has found a technical application in the explosive argon flash lamp (Winning and Edgerton 1952). Table 4 shows the calculated dependence of U, p_2 and T_2 on p_0 (Döring and Burkhardt 1944). As exact data are lacking the simplifying assumption has been made that U and T_2 depend on T_0 and p_2/p_0 only.

Table 4. Influence of the Air Pressure on the Shock Wave (Johansson and Ljungberg 1953). Explosive: PETN + TNT

p_0 (atm)	p_2 (atm)	U (m sec^{-1})	T_2 (°K)
10	4650	6500	8200
1	1040	9500	13100
0·1	190	13000	23400
0·01	32·5	17000	33200
0·001	4·8	20500	41000

Since the energy-consuming ionization, whose dependence upon pressure has been neglected, is favoured by low pressure, the values T_2 indicated in table 4 for $p_0 < 1$ atm are probably too high.

Instead of plane wave fronts, we will consider now curved wave fronts. If a curved detonation front touches a curved charge surface, essentially the same considerations as above may be applied, e.g. in the case of a spherical charge ignited in its centre. If the detonation front hits the surface at an angle $\phi \neq 0$ the shock wave is weaker, but still of the same order as in the case $\phi = 0$; calculation by means of a Prandtl–Meyer expansion fan (Döring and Burkhardt 1944) is possible.

During its spatial propagation the shock wave is attenuated. Under simple geometrical conditions, e.g. in the case of spherical symmetry, one can calculate the whole expansion process (Wecken and Mücke 1950, 1953, Berry *et al.* 1955). In addition experimental investigations have furnished empirical formulae describing the spherical shock propagation (Ericsson 1955). For charges of less regular shape the shock strength diminishes most quickly in the directions of greatest convexity, especially at the edges; hence the cross-formed luminosity produced by a detonating quad (Muraour 1953). If the charge surface is lined by a metal sheet, the product gas at first accelerates the metal which then produces a shock wave in the surrounding gas (Christian *et al.* 1955, Deal 1954). If the metal plate quickly gets a high constant velocity u one can measure at the same time u and U and thus find the Hugoniot curve of the gas. Such investigations have been made for air (Deal 1954), argon (Christian *et al.* 1955), and nitrogen (Christian *et al.* 1955). For an unlined charge the boundary of the product gas is quickly deformed owing to an instability (Wecken and Mücke 1953, Markstein 1957), thus making the u_2 measurement inexact.

6.2.2. *The luminosity of shock waves.*

A strong shock wave in air, comparable with a detonation wave, consists of a steep pressure rise front on the low pressure side and of a broader transition zone (Döring 1949). In the shock front only the translational and rotational energy of the gas molecules adjust to a new state corresponding to the shock velocity U; the other degrees of freedom (oscillation, excitation, dissociation, ionization) are unaltered; they adjust to the new equilibrium in the transition zone. Here

p, v and T_{transl} change continuously. p and v are linearly connected by the equation of state line $(v_0 - v)U^2 = v_0^2(p - p_0)$ (cf. equation (67)).

The luminosity depends on the ionization and therefore does not reach its full intensity before the end of the transition zone.

The radiation of ionized gases shows a continuous spectrum and is essentially caused by retardation or capture of free electrons in the Coulomb field of ions (Sauter 1949). The spectral energy distribution (Michel-Lévy *et al.* 1941) in the case of an optically thin layer deviates decidedly from black-body radiation. Burkhardt (1951) investigated theoretically the radiation as a function of temperature, density, layer thickness and wavelength. His results for a special detonation process have been compared with experiments. The first intensity maximum measured was explained satisfactorily by the shock wave theory. A second maximum of longer duration seems to be caused by chemical reactions between product gases and air.

At the encounter of two detonation shock waves in argon Vassy (Michel-Lévy *et al.* 1941) found an intense luminosity with an energy distribution corresponding to a black-body radiation of 35 000°K. Determination of the absolute intensity was not possible, but supposedly the optical thickness was sufficient to make colour temperature equal to real temperature (Sauter 1949).

6.2.3. *Shock waves from nuclear explosions.*

An atomic explosion in air (*Atomic Weapons* 1950, *Nuclear Explosions* 1956) produces a strong shock wave which in some respects differs from the shock wave generated by a chemical explosive. A fairly good description of the free propagating shock wave and the pressure distribution is given by the point blast theory (Taylor 1950) from which we gather the following data.

In a perfect gas ($p = p_0$, $\rho = \rho_0$) at the point $r = 0$ a finite amount E of energy is suddenly released at the time $t = 0$. For $t > 0$ this energy is spread over a sphere $r < R(t)$ limited by a spherical shock wave of radius

$$R(t) = S(\gamma) t^{2/5} E^{1/5} \rho_0^{-1/5} \qquad \ldots\ldots (71)$$

and of velocity $U = 0{\cdot}4R/t$. For $\gamma = 1{\cdot}4$, $S = 1{\cdot}03$, and the shock pressure $p_s = 0{\cdot}15\delta\, ER^{-3}$. At $t = \text{const.}$, p increases monotonically with r from $p = 0{\cdot}373\, p_s$ at $r = 0$ to $p = p_s$ at $r = R(t)$, while T decreases monotonically. This theory holds only as far as $p_s \gg p_0$.

In a test explosion (Taylor 1950) equation (71) was indeed found to be approximately valid for 27 km sec$^{-1} > U > 1{\cdot}2$ km sec^{-1}, and it probably holds, according to the scaling law, for all nuclear explosions within the same limits. One defines $E = 8{\cdot}4 \times 10^{20}$ erg for a 'nominal atomic explosion', $E = 8{\cdot}4 \times 10^{23}$ erg for a 'nominal high-yield explosion'. The shock wave develops as soon as the expansion of the fire ball proceeds faster by molecule shocks than by radiation; for a nominal atomic explosion this is the case at $t = 0{\cdot}1$ msec, $R = 13$ m, $T_s = 300\,000$°K, $p_s \simeq 25\,000$ atm, $U = 85$ km sec^{-1}. A shock wave of greater strength in air does not seem to be possible. Later on the radiation front remains behind the shock front and T_s, p_s and U decrease monotonically with increasing t and R.

For temperatures from 50 000 to 300 000°K the radiation has a considerable influence both on the structure of the shock front and on the energy balance of

the shock wave; an investigation of these effects has not been published. For $T_s < 50\,000°K$ the reaction of the radiation on the shock wave itself is unimportant. For $T_s > 2000°K$ the layer between shock and radiation front is still optically thick and the shock sphere radiates outwards nearly as a black body with $T = T_s$. At $T_s \simeq 2000°K$ the shock front 'breaks away', i.e. the shocked air becomes transparent and makes the hot core visible.

References

Allen, K. W., Bodin, H. A. B., Curran, S. C., and Fitch, R. A., 1957, III Congr. Fenomeni D'Ionizzazione Nei Gas (Milan: Venezia Rendiconti).

Aller, L. H., 1953, *Astrophysics* (New York: Ronald Press).

Anderson, J. A., 1920, *Astrophys. J.*, **51**, 37.

Anderson, J. A., and Smith, S., 1926, *Astrophys. J.*, **64**, 295.

Anderson, O. A., Becker, W. R., Colgate, S. A., Ilse, J., and Pyle, R. V., 1957, III Congr. Fenomeni D'Ionizzazione Nei Gas (Milan: Venezia Rendiconti).

Bartels, H., 1949 a, *Z. Phys.*, **125**, 598; 1949 b, *Ibid.*, **126**, 108; 1950 a, *Ibid.*, **127**, 243; 1950 b, *Ibid.*, **128**, 546; 1956, Plasma Conference, Leipzig.

Bartels, H., and Eiselt, B., 1950, *Optik, Stuttgart*, **6**, 56.

Bates, D. R., and Damgaard, A., 1949, *Phil. Trans. Roy. Soc.* A, **242**, 101.

Bennett, W. H., 1934, *Phys. Rev.*, **45**, 890.

Berry, F. J., Butler, D. S., and Holt, M., 1955, *Proc. Roy. Soc.* A, **227**, 258.

Biermann, L., 1950, *Landolt-Börnstein Tables*, Vol. I, Part 1, p. 260 (Berlin: Springer).

Biermann, L., and Schlüter, A., 1957, *Z. Naturf.*, **12a**, 805.

Bogen, P., 1957, *Z. Phys.*, **149**, 62.

Bötticher, W., 1957, *Z. Phys.*, **150**, 336.

Bowden, F. P., and Yoffe, A., 1952, *The Initiation of Explosions* (Cambridge: University Press).

Burhorn, Fr., 1955, *Z. Phys.*, **140**, 440; 1956, *Phys. Verh.*, **7**, 198.

Burhorn, Fr., and Maecker, H., 1951, *Z. Phys.*, **129**, 369.

Burkhardt, G., 1940, *Z. Phys.*, **115**, 592; 1951, *St. Louis Lab. News Rapp.*, **14/51**, 309.

Busz, G., and Finkelnburg, W., 1953, *Naturwissenschaften*, **40**, 550; 1954, *Z. Phys.*, **138**, 212; 1956, *Ibid.*, **146**, 655.

Butler, D. S., 1955, *Symposium on Blast and Shock Waves, Fort Halstead* (*A.R.D.E. Memo* (*B*), 19/56).

Campbell, A. W., Malin, M. E., Boyd, T. J., and Hull, J. A., 1955, O.N.R. Symposium, Washington.

Chabai, A. J., and Emrich, R. J., 1955, *J. Appl. Phys.*, **26**, 779 (L).

Chandrasekhar, S., 1945, *Astrophys. J.*, **102**, 401.

Christian, R. H., Duff, R. E., and Yarger, F. L., 1955, *J. Chem. Phys.*, **23**, 2042, 2045.

Compton, K. T., 1923, *Phys. Rev.*, **21**, 266.

Conn, W. M., 1955, *Z. angew. Phys.*, **7**, 539.

Cook, M. A., 1947, *J. Chem. Phys.*, **15**, 518.

Deal, W. E., 1954, *Phys. Rev.*, **94**, 811.

Demougin, P., 1930, *Mémor. Artill. Fr.*, **9**, 961.

Ditchburn, R. W., Jutsum, P. J., and Marr, G. V., 1953, *Proc. Roy. Soc.* A, **219**, 89; 1954, *Ibid.*, **224**, 83; see also Marr, G. V., 1955, *Proc. Phys. Soc.* A, **68**, 544.

Dolgow, G. G., and Mandelstam, C. L., 1953, *J. Exp. Theor. Phys., Moscow*, **24b**, 691.

Döring, G., 1957, *Dissertation*, University of Kiel.

Döring, W., 1949, *Ann. Phys., Lpz.*, **5**, 133; 1943, *Ibid.*, **43**, 421.

Döring, W., and Burkhardt, G., 1939, *Luftfahrtforschung* (Berlin: Adlerhof); 1944, *ZWB Forschungsberichte*, **222**, 64.

Drawin, H. W., 1956, *Z. Phys.*, **146**, 295.

Duff, R. E., and Houston, E., 1955, *J. Chem. Phys.*, **23**, 1268.

ECKER, G., 1951, *Z. Phys.*, **130**, 585.
EGGERT, J., 1919, *Z. Phys.*, **20**, 570.
EISELT, B., 1952, *Z. Phys.*, **132**, 54.
ELENBAAS, W., 1951, *The High Pressure Mercury Vapour Discharge* (Amsterdam : North-Holland), p. 39.
VOM ENDE, 1958, *Diplomarbeit*, Kiel University.
VON ENGEL, A., 1955, *Ionized Gases* (Oxford : Clarendon Press), p. 239.
VON ENGEL, A., and STEENBECK, M., 1931, *Wiss. Veröff. Siemens*, **10**, 155 ; 1933, *Ibid.*, **12**, 74.
MCENHILL, J. J., 1955, *Symposium on Blast and Shock Waves, Fort Halstead* (*A.R.D.E. Memo* (*B*), 19/56).
ERICSSON, U., 1955, *Appl. Sci. Res.*, **a5**, 309.
EULER, J., 1951, *Z. angew. Phys.*, **3**, 260 ; 1953 a, *Ibid.*, **5**, 64 ; 1953 b, *Ann. Phys., Lpz.*, **11**, 203.
EVANS II, G. W., FAULKNER, F. D., LOCKHART, B. J., and PERRY, C. L., 1956, *J. Appl. Phys.*, **27**, 1345.
EYRING, H., POWELL, R. E., DUFFEY, G. H., and PARLIN, R. B., 1949, *Chem. Rev.*, **45**, 69.
FAST, J. D., 1947, *Philips Res. Rep.*, **2**, 382.
FINKELNBURG, W., 1948 a, *Hochstromkohlebogen* (Berlin : Springer) ; 1948 b, *Phys. Rev.*, **74**, 1475 ; 1948 c, *Ibid.*, **74**, 222 ; 1956, *Handbuch der Physik* (Ed. H. Flügge), **22**, 305 (Berlin : Springer).
FINKELSTEIN, R. J., and GAMOW, G., 1947, Theory of the Detonation Process, *Navord Report* 90-46.
FISCHER, H., 1955, *Phys. Verhandl., Mosbach*, **6**, 177 ; 1956 a, *Nucleonics*, **14**, 42 ; 1956 b, *Ibid.*, **14**, 36.
FONBERG, Z., 1951, *J. Chem. Phys.*, **19**, 383.
FOWLER, R. G., ATKINSON, W. R., COMPTON, W. D., and LEE, R. J., 1952, *Phys. Rev.*, **88**, 137.
FOWLER, R. G., ATKINSON, W. R., and MARKS, L. W., 1952, *Phys. Rev.*, **87**, 966.
FOWLER, R. H., 1925, *Mon. Not. R. Astr. Soc.*, **85**, 970 ; 1936, *Statistical Mechanics* (Cambridge : University Press).
FOWLER, R. H., and MILNE, E. A., 1923, *Mon. Not. R. Astr. Soc.*, **83**, 403 ; 1924, *Ibid.*, **84**, 499.
FOX, J. G., 1945, *U.S. Navord Report* 200-45.
FREEMAN, N. C., 1955, *Proc. Roy. Soc.* A, **228**, 341.
FREIWALD, H., and UDE, H., 1957, *Z. Elektrochem.*, **61**, 663.
FUCHS, R., 1951, *Z. Phys.*, **130**, 69.
FÜNFER, E., 1958, *Z. angew. Phys.*, in the press.
FÜNFER, E., and MÜLLER, W., 1956, *2nd Congress of High-speed Photography*, 1954, *Paris* (Paris : Dunod).
GAMOW, G., and CRITCHFIELD, C. L., 1949, *Theory of Atomic Nuclei and Nuclear Energy Sources* (Oxford : Clarendon Press), p. 264.
GARSTANG, R. H., 1950, *Mon. Not. R. Astr. Soc.*, **110**, 613 ; 1951, *Ibid.*, **111**, 115.
GAYDON, A. G., and PENNEY, W. G., 1945, *Proc. Roy. Soc.* A, **183**, 374.
GERDIN, H., and LOTZ, A., 1922, *Wiss. Veröff Siemens*, **2**, 489 ; 1923, *Z. techn. Phys.*, **4**, 157 ; 1924, *Ibid.*, **5**, 515 ; 1935, *Wiss. Veröff Siemens*, **14**, 25.
GIBSON, F. C., BOWSER, M., SUMMERS, C. R., SCOTT, F., and MASON, C. M., 1957, *Bull. Amer. Phys. Soc.*, **2**, 35.
GÖING, W., 1951, *Z. Phys.*, **131**, 603.
GRIEM, H., 1954, *Z. Phys.*, **137**, 280.
GUDERLEY, G., 1942, *Luftfahrtforsch.*, **19**, 302.
HÄFELE, W., 1956, *Z. Naturf.*, **11a**, 183.
VAN DER HELD, E. F. M., 1931, *Z. Phys.*, **70**, 508.
HILL, R., and PACK, D. C., 1947, *Proc. Roy. Soc.* A, **191**, 524.
HOLTSMARK, J., 1919 a, *Ann. Phys., Lpz.*, **58**, 577 ; 1919 b, *Z. Phys.*, **20**, 162 ; 1924, *Ibid.*, **25**, 73.

TER HORST, D. TH. J., and RUTGERS, G. A. W., 1953, *Physica*, **19**, 565.
INGLIS, D. R., and TELLER, E., 1939, *Astrophys. J.*, **90**, 439.
JOHANSSON, C. H., and S. LJUNGBERG, 1953, *Ark. Fys.*, **6**, 369.
JÜRGENS, G., 1952, *Z. Phys.*, **134**, 21.
KENTY, C., 1941, *Phys. Rev.*, **60**, 66.
KOCH, B., 1953, *C.R. Acad. Sci., Paris*, **236**, 661.
KOCH, O., 1949, *Z. Phys.*, **126**, 507.
KOLB, A. C., 1956, *Ph.D. Thesis*, University of Michigan ; 1957, *Lockheed Magnetohydrodynamics Symposium* (Stanford : University Press ; Oxford : University Press).
KRUITHOF, A. M., 1943 a, *Physica*, **10**, 493 ; 1943 b, *Dissertation*, Utrecht, p. 18.
KRUITHOF, A. M., and SMIT, J. A., 1944, *Physica*, **11**, 129.
KUHN, H., 1934, *Phil. Mag.*, **18**, 987.
LARCHÉ, K., 1952, *Z. Phys.*, **132**, 544 ; 1951, *Z. Elektrotech., Stuttgart*, **72**, 427.
LARENZ, R. W., and BARTELS, H., 1950, *Naturwissenschaften*, **37**, 164 ; 1951, *Z. Phys.*, **129**, 327, 343.
LEBEAU, P., and TROMBE, F., 1950, *Les Hautes Températures et Leurs Utilisations en Chimie* (Paris : Masson), **1**, p. 151.
LENZ, W., 1933, *Z. Phys.*, **80**, 423.
LIN, S.-C., RESLER, E. L., and KANTROWITZ, A., 1955, *J. Appl. Phys.*, **26**, 95.
LINDHOLM, E., 1941, *Ark. Mat. Astr. Fys.*, **28B**, 3 ; 1942, *Dissertation*, University of Uppsala ; 1945, *Ark. Mat. Astr. Fys.*, **32A**, 17.
LÖCHTE-HOLTGREVEN, W., 1951, *Conference sur la Combustion, la Detonation, et les Ondes de Choc*, St. Louis Memoire Nr 14m/51 ; 1955, *Appl. Sci. Res.*, **B5**, 182 ; 1956, Plasma Conference, Leipzig ; 1957 a, *Proceedings of the Colloquium on Spectroscopy*, Vol. 6, p. 111 (London : Pergamon Press) ; 1957 b, International Conference on Ionization, Venice.
LOCHTE-HOLTGREVEN, W., and NISSEN, W., 1952, *Z. Phys.*, **133**, 124.
LORENTZ, H. A., 1905, *Verh. Akad. Wet. Amst.*, **14**, 518, 577.
MAECKER, H., 1951, *Z. Phys.*, **129**, 108 ; 1953, *Ibid.*, **136**, 119 ; 1955, *Ibid.*, **141**, 198 ; 1956, *Ann. Phys., Lpz.*, **18**, 441 ; 1956, *Z. Naturf.*, **11a**, 457.
MAECKER, H., and PETERS, TH., 1954, *Z. Phys.*, **139**, 448.
MAECKER, H., PETERS, TH., and SCHENK, H., 1955, *Z. Phys.*, **140**, 119.
MANDELSTAM, S., 1956 a, *Exp. Techn. Phys.*, **3**, 50 ; 1956 b, Plasma Conference, Leipzig.
MANNKOPFF, R., 1932, *Z. Phys.*, **76**, 396.
MARGENAU, H., 1932, *Phys. Rev.*, **40**, 387 ; 1957, *Yale University Report*.
MARKSTEIN, G. H., 1957, *J. Aero Sci.*, **24**, 238.
MASSEY, H. S. W., and MOISEIWITSCH, B. L., 1951, *Proc. Roy. Soc.* A, **205**, 483.
MCPHERSON, H. G., 1940, *J. Opt. Soc. Amer.*, **30**, 189.
MEEK, J. M., and CRAGGS, J. D., 1953, *Electrical Breakdown of Gases* (Oxford : University Press).
MICHEL-LÉVY, A., MURAOUX, H., and VASSY, E., 1941, *Rev. Opt.* (*Théor. Instrum.*), **20**, 149.
MINKOWSKI, R., 1930, *Z. Phys.*, **63**, 188.
MOWER, L., 1953, *Phys. Rev.*, **89**, 947.
MÜLLER, W., 1957, *Z. Phys.*, **149**, 397.
MURAOUR, H., 1953, *C.R. Acad. Sci., Paris*, **237**, 514.
MURGAI, M. P., 1953, *J. Chem. Phys.*, **21**, 403.
NISSEN, W., 1953, *Diplomarbeit*, University of Kiel ; 1954, *Z. Phys.*, **139**, 638.
ORNSTEIN, L. S., and BRINKMAN, H., 1934, *Physica*, **1**, 797.
PALMER, H. B., 1956, *J. Appl. Phys.*, **27**, 1105.
PATERSON, S., 1955, *5th Symposium on Combustion*, p. 672 (New York : Reinhold).
PAYNE, C. H., 1925, "Stellar Atmosphere", *Harv. Obs. Monogr.*, **1**, 132.
PENNEY, W. G., 1950, *Proc. Roy. Soc.* A, **204**, 1.
PERRY, R. W., and KANTROWITZ, A., 1951, *J. Appl. Phys.*, **22**, 878.
PESCHEK, H. E., ROSE, P. H., GLICK, H. S., KANE, A., and KANTROWITZ, A., 1955, *J. Appl. Phys.*, **26**, 83.

PETERS, TH., 1953, *Z. Phys.*, **135**, 573 ; 1954 a, *Phys. Verh.*, **5**, 163 ; 1954 b, *Naturwissenschaften*, **41**, 571.
PORITSKY, H., and SUITS, C. G., 1935, *Physica*, **6**, 190.
RESLER, E. L., LIN, S.-C., and KANTROWITZ, A., 1952, *J. Appl. Phys.*, **23**, 1390.
REYNOLDS, P., and CRAGGS, J. D., 1942, *Phil. Mag.*, **43**, 258.
RÖSSLER, F., and SCHALL, R., 1953, *St. Louis Lab. Report.*
SÄNGER, E., 1954, 27th Congr. Chim. Industr., Brussels.
SAHA, M. N., 1920, *Phil. Mag.*, **40**, 472 ; 1921, *Z. Phys.*, **6**, 40.
SAUTER, F., 1949, *St. Louis Lab. Report*, 17/49.
SAVITT, J., and STRESAU, R. H. F., 1954, *J. Appl. Phys.*, **25**, 89.
SCHALL, R., and THOMER, G., 1951, *Z. angew. Phys.*, **3**, 41.
SCHARDIN, H., 1954, *Commun. Pure Appl. Math.*, **7**, 223.
SCHLÜTER, A., 1957, *Z. Naturf.*, **12a**, 822.
SCHMALJOHANN, 1936, *Staatsexamensarbeit*, Kiel University.
SCHMIDT, P., 1955, Patentschrift München (Sch. 15599 VIII c/21g).
SCHMITZ, G., 1949, *Z. Phys.*, **126**, 1.
SCHUTTEVAER, J. W., DE BONT, M. J., and VAN DEN BROEK, TH. H., 1943, *Physica*, **10**, 546.
SHREFFLER, R. G., and CHRISTIAN, R. H., 1954, *J. Appl. Phys.*, **25**, 324.
SMIT-MIESSEN, M. M., and SPIER, J. L., 1942, *Physica*, **9**, 193.
SPERLING, J., 1950, *Z. Phys.*, **128**, 209.
SPIER, J. L., and SMIT, J. A., 1942, *Physica*, **9**, 597.
SPIER, J. L., and SMIT-MIESSEN, M. M., 1942, *Physica*, **9**, 422.
SPITZER, L., JR., 1956, *Physics of Fully Ionized Gases* (New York : Interscience Tracts, No. 3).
STEINHAUS, D. W., CROSSWHITE, H. M., and DIEKE, G. H., 1953, *Spectrochim. Acta*, **5**, 436.
STETTBACHER, A., 1948, *Spreng und Schießstoffe*, p. 16 (Zürich : Rascher).
SUITS, C. G., 1935, *Physica*, **6**, 315.
TAYLOR, G. I., 1950, *Proc. Roy. Soc.* A, **201**, 159, 175.
TAYLOR, J., 1952, *Detonation in Condensed Explosives* (Oxford : University Press).
TAYLOR, R. J., 1957, *Proc. Phys. Soc.* B, **70**, 31.
THONEMANN, P. C., BUTT, E. P., CARRUTHERS, R., DELLIS, A. N., FRY, D. W., GIBSON, A., HARDING, G. N., LEES, D. J., MCWHIRTER, R. W. P., PEASE, R. S., RAMSDEN, S. A., and WARD, S., 1958, *Nature, Lond.*, **181**, 217. (See also the following articles in the same issue, pp. 221–233.)
THONEMANN, P. C., and COWHIG, W. T., 1950, *Nature, Lond.*, **166**, 903.
TONKS, L., 1937, *Trans. Electrochem. Soc.*, **72**, 167.
TONKS, L., and ALLIS, W., 1939, *Phys. Rev.*, **56**, 360.
TURNER, E. B., 1956, *Ph.D. Thesis*, Michigan University.
UNSÖLD, A., 1938, *Ann. Phys., Lpz.*, (5) **33**, 607 ; 1948, *Z. Astrophys.*, **24**, 355 ; 1955, *Physik der Sternatmosphären* (Berlin : Springer), pp. 288, 306.
VERWEY, S., 1936, *Dissertation*, University of Amsterdam.
WECKEN, F., and MÜCKE, L., 1950, *St. Louis Lab. Report*, 8/50 ; 1953, *Ibid.*, 1/53.
WEIBULL, W., 1947, *Ark. Mat. Astr. Fys.*, **34b**, 1.
WEISS, R., 1954, *Z. Phys.*, **138**, 170.
WEISSKOPF, V., 1932, *Z. Phys.*, **75**, 287.
WIENECKE, R., 1955, *Z. Phys.*, **143**, 128 ; 1956, *Ibid.*, **146**, 39 ; 1958, *Ibid.*, **150**, 231.
WINNING, C. H., and EDGERTON, H. E., 1952, *J. Soc. Mot. Pict. Engrs*, **59**, 178.
WRANA, J., 1939, *Arch. Elektrotechn.*, **33**, 656.
WULFF, H., 1958, *Z. Phys.*, **150**, in the press.